PIER

PUEY UNGPHAKORN INSTITUTE
FOR ECONOMIC RESEARCH

Farmers and Pixels: Toward Sustainable Agricultural Finance with Space Technology

by

Sommarat Chantarat, Preesan Rakwatin and Chutatong Charumilind

October 2017

Discussion Paper

No. 75

Farmers and Pixels: Toward Sustainable Agricultural Finance with Space Technology[†]

Sommarat Chantarat
Puey Ungphakorn Institute for Economic Research

Preesan Rakwatin
Digital Economy Promotion Agency

Chutatong Charumilind
Insurance Premium Rating Bureau

October 26, 2017

Abstract

This paper explores promises of satellite technology in creating high-quality agricultural risk information necessary for unlocking market inefficiencies that have precluded sustainable development of insurance markets and overall risk management in agricultural sector, where uninsured risk remains a leading impediment of economic development. Using pixel-level, high resolution, high frequency and longitudinal satellite data together with a combination of geographical information system (GIS) data, administrative and household-level agricultural data, this paper answers three questions: (1) Can satellite data be used to generate high-quality risk information for Thai rice farmers? (2) How might the satellite-based risk information be used to crowd in sustainable markets for agricultural finance? And (3) What are potential economic impacts of having high quality agricultural data on farmers, agricultural banks and government? After illuminating the potential values of investing in high-quality agricultural data, this paper also discusses key challenges and ways forward in bringing this research into real action to enhance financial stability of farmers, financial system and government.

JEL Codes: D81, G22, Q12, Q14, Q18

Keywords: Satellite technology, Agricultural risk, Agricultural insurance, Risk-contingent credit, Rice production, Thailand

---

[†] The authors can be contacted at sommarac@bot.or.th; preesan.ra@depa.or.th; chutatong@tgia.org. We thank Geo-informatics and Space Technology Development Agency (GISTDA), Office of Agricultural Promotion, Office of Agricultural Economics and Bank of Agriculture and Agricultural Cooperatives (BAAC) for providing data, and Nipon Poapongsakorn and Chularat Tanprasert for their valuable comments and suggestions. Panu Nuangjumnong provides excellent research assistance. The views expressed in this study are our own and do not represent those of the Bank of Thailand, Digital Economy Promotion Agency or Insurance Premium Rating Bureau. All errors are our own.

Introduction

Sustainable agricultural finance is critical for economic development of farming households. With less than 1% insured, recent research shows that Thai farmers still could not smooth their consumption against increasing covariate shocks, most of which are driven by climate and directly affect agricultural production. Most farmers rely on income diversification and risk-sharing mechanisms (which appear ineffective in dealing with most agricultural shocks that often affect wider economy) and government's disaster relief (which are largely inadequate and delayed, albeit increasing public spending every year). Large agricultural shocks, like the current drought, thus not only directly result in sharp drop of farmers' consumption and rising debt accumulation and non-performing loans of agricultural credit portfolios, but also indirectly affect farmers' long-term prospects by reducing their investment incentives and ability. Uninsured shocks thus reduce credit market deepening by reducing both credit demand and supply – leaving the government supported Bank of Agriculture and Agricultural Cooperatives (BAAC) the only player in the market with significant portfolio losses and increasing budget burden to government.

Sustainable agricultural finance requires healthy risk market, development of which could face many challenges, most of which can only be resolved using high-quality risk information (accurate, transparent, long history, fast and low cost). Sustainable insurance market relies on capacity of insurers to transparently monitor and obtain true risk information of insured farmers at low cost to ensure of effective risk-based pricing and in near real time so to ensure value to farmers. This can resolve the common information asymmetry, which could otherwise result in market failure when insurers are forced to charge prohibitive premium to protect themselves from the potential moral hazard and adverse selection problems. But resolving information asymmetry rather comes with extremely high transaction cost (which again could result in prohibitive premium per contract given that plots are small but remote and scattered in Thailand) and long delay (which could further reduce the value of insurance to farmers and market demand). Given the covariate nature of agricultural risk, sustainable agricultural insurance also relies on capacity of local insurers to cost-effectively transfer extreme risk to international market, which again could be possible with transparent insured risk information.

Agricultural insurance in Thailand is still in early stage of development with a well-established nationwide program for rice farmers distributed by BAAC since 2011, but still face a great deal of challenges. The program is more than 50% subsidized by government and relies on declaration of disaster areas and farm-level loss verification by local government (which is already in place for public disaster relief) to determine insurance payouts, and thus could reduce transaction cost accordingly. To the extent that government's loss verification could be less transparent to insurers (e.g., could be in favor of farmers especially during election period) and with the lack of transparent historical risk profile of insured farmers, the program thus still largely suffers information asymmetry, resulting in extremely high premium. The program could further suffer from long delay in loss verification affecting insurance demand. And so overall, less than 1% of rice growing areas has been insured to date despite the extremely high public subsidization. Sustainable agricultural insurance in Thailand will need better risk information.

Satellite technology has great promise in providing necessary risk information for unlocking challenges underpinning sustainable development of agricultural insurance and credit markets. Satellite data are derived from signals and imagery received from earth observatory satellites. These data exist for quite a period of time but recent revolution in data analytics makes it possible for researchers, policy makers and businesses to take advantage of their key features. They can thus best reflect geographical conditions at granular level with high frequency and in near real time. And more importantly, they could be obtained at low cost.

Satellite data could thus provide insurers with good and transparent proxy for risk information of representative farmers at very localized area in near real time and with long historical profile. Thus index-based agricultural insurance products designed to trigger payouts in each localized area based on these data could hold great potential in resolving various market failures in agricultural insurance market allowing for risk-based pricing, cost effective risk transfer and timely payouts to farmers with low transaction cost. These potentials however rest upon the extent to which satellite data can reflect true agricultural risk faced by insured farmers in each localized insured area and how farmers understand and trust this technology. These promises rely on the potential that satellite data could be transparent, long history, near real time, relatively low cost. But like other data, still contains errors, to what extent can they accurately reflect actual risk faced by farmers?

The goal of this paper is to explore the promises of satellite technology and GIS data in creating high-quality risk information for farmers so to unlock inefficiencies that plague the development of agricultural insurance and credit markets and overall risk management program of the country. Using pixel-level data from two satellites with moderate-high resolution, high frequency and long historical series combined with a combination of geographical information system (GIS) data, administrative data and household-level agricultural data, this paper answers three research questions: (1) Can satellite data be used to generate necessary risk information for Thai rice farmers? (2) What are potential values of satellite-based risk information? How the satellite-based risk information be used to create sustainable insurance and credit markets? And (3) What are potential economic impacts on farmers, agricultural banks and government? How can this be compared with existing scheme?

The paper starts by reviewing currently available agricultural data and reflects on nature and vulnerability of rice production and on current inefficiencies that impede development of sustainable agricultural risk management. Section 2 reviews recent development in satellite technology and their applications in agriculture emphasizing on the products, current applications and initiatives in Thailand. Section 3 then explores the use of satellite data to construct risk data and explore their performance using household-level data. The section also illustrates various kinds of risk information that can be reviewed using satellite data. Using this newly constructed satellite-based risk data, section 4 then designs various insurance contracts for farmers, groups of farmers and BAAC to better manage agricultural risk, explores risk-based pricing and diversification and explores public-private partnership arrangements, where support from government can crowd in healthy insurance market and risk management incentives of farmers. Section 5 conducts simulation-based assessment of potential impacts of these market arrangements, illuminating the potential benefit in enhancing financial stability of farmers, BAAC and government. Section 6 concludes with lively discussion on ways forward, key challenges and the required institutional and policy arrangements that might be needed in bringing the promises of satellite data illuminated in this research into real action toward sustainable development agricultural finance and livelihoods.

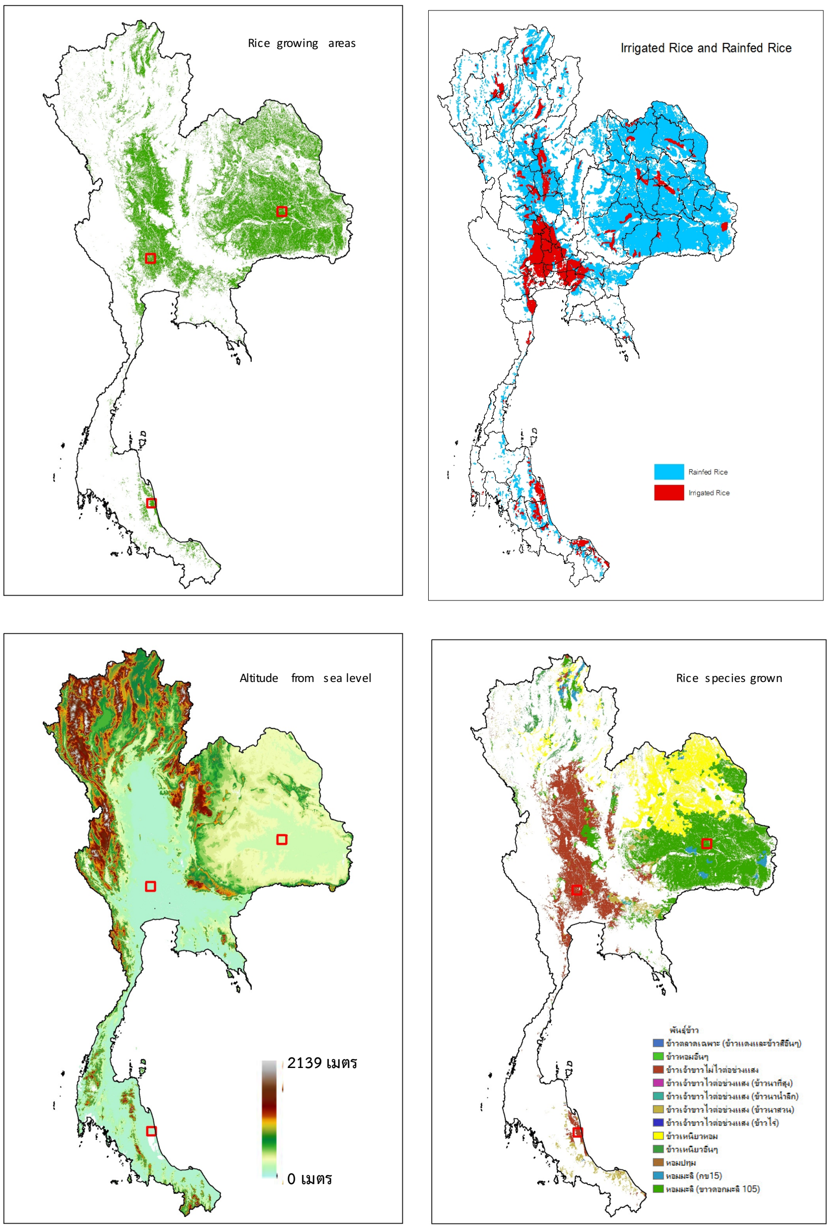
1. Current agricultural risk data, rice production and inefficiencies in agricultural finance

This paper begins by exploring the wealth of available agricultural data in Thailand to understand the rice production system, the values and pitfalls of existing data and the current inefficiencies in credit and insurance market, which are due mainly to information asymmetry. These data would then be used either to combine with satellite data in creating better risk information or to verify the predictive performance of the satellite data.

Current agricultural data

We first turn to available GIS information that can illuminate the nation's rice production. Figure 1 shows rice production areas constructed from both imagery analysis and extended ground trothing, irrigation map, altitude and rice species.

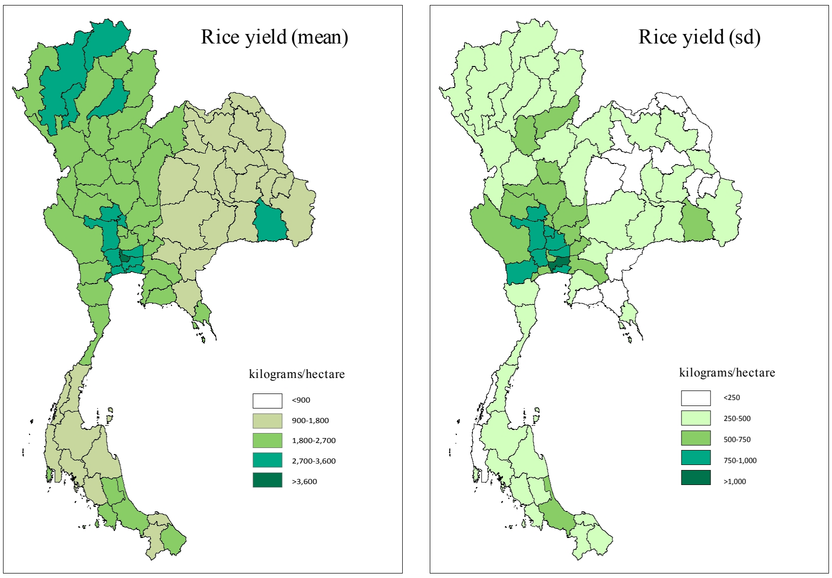
Figure 1: Available GIS data on Thailand's rice production



There are great variations in rice production across more than 70 million rai growing areas, with less than half of area having irrigation. Majority of rice areas are low land with some exception in the north and relatively higher flat plan in the northeast. Species also vary geographically.

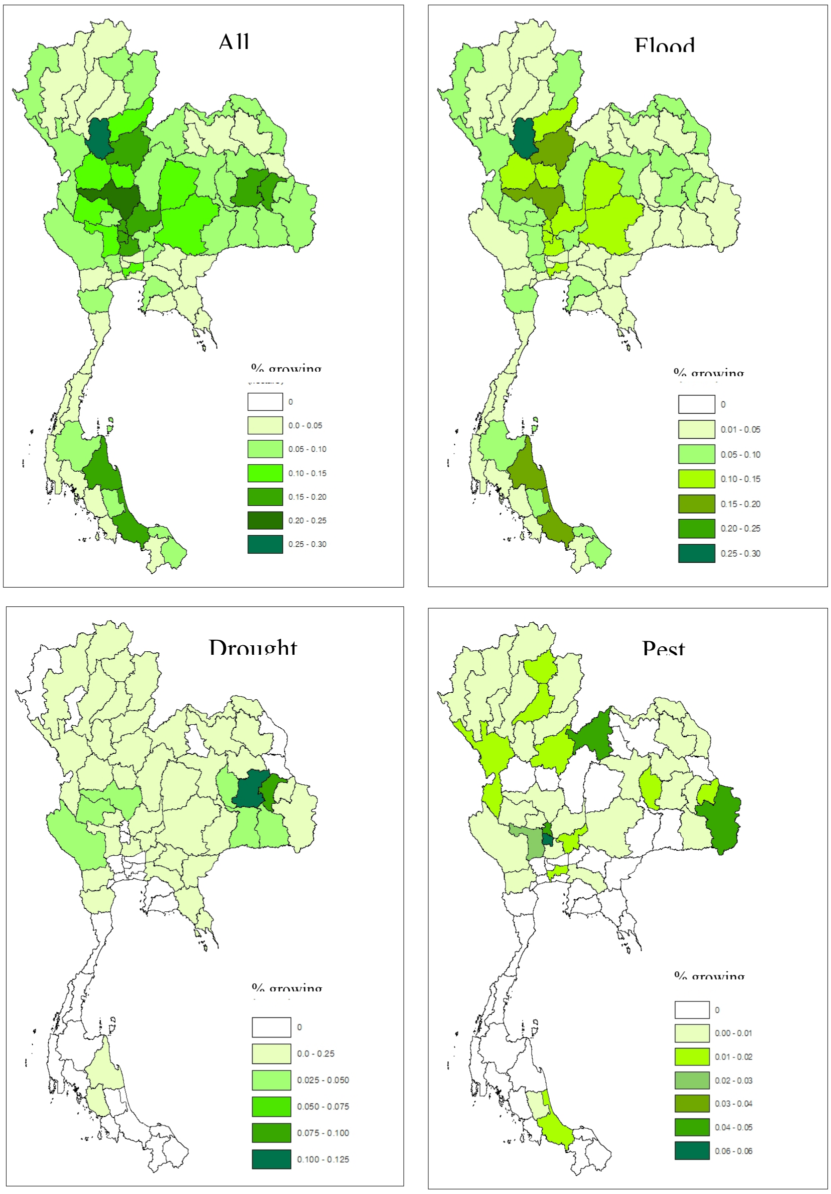
Figure 2 further shows that rice yield also varies geographically. Yield data has been collected with more than 30 years of historical profile but at provincial level from the Office of Agricultural Economics. So less is known about what happen within province or how risk exposure and production could vary within province each year. Amphoe-level yield data just recently become available from 2554-2557 making it more possible to understand agricultural risk profile at more micro level. Unfortunately, the series is not yet long enough. Moreover, these yield data are based on field data collection and reports from local agricultural offices, which could take great amount of time (at least a year lag) until the arrival of the new season rice. So though they review great geographical differences in term of yield, the data would be less useful in term of production monitoring, disaster detection and intervention and/or insurance, which requires more near-real time measures.

Figure 2: Rice yield by Amphoe (30 years until 2557)



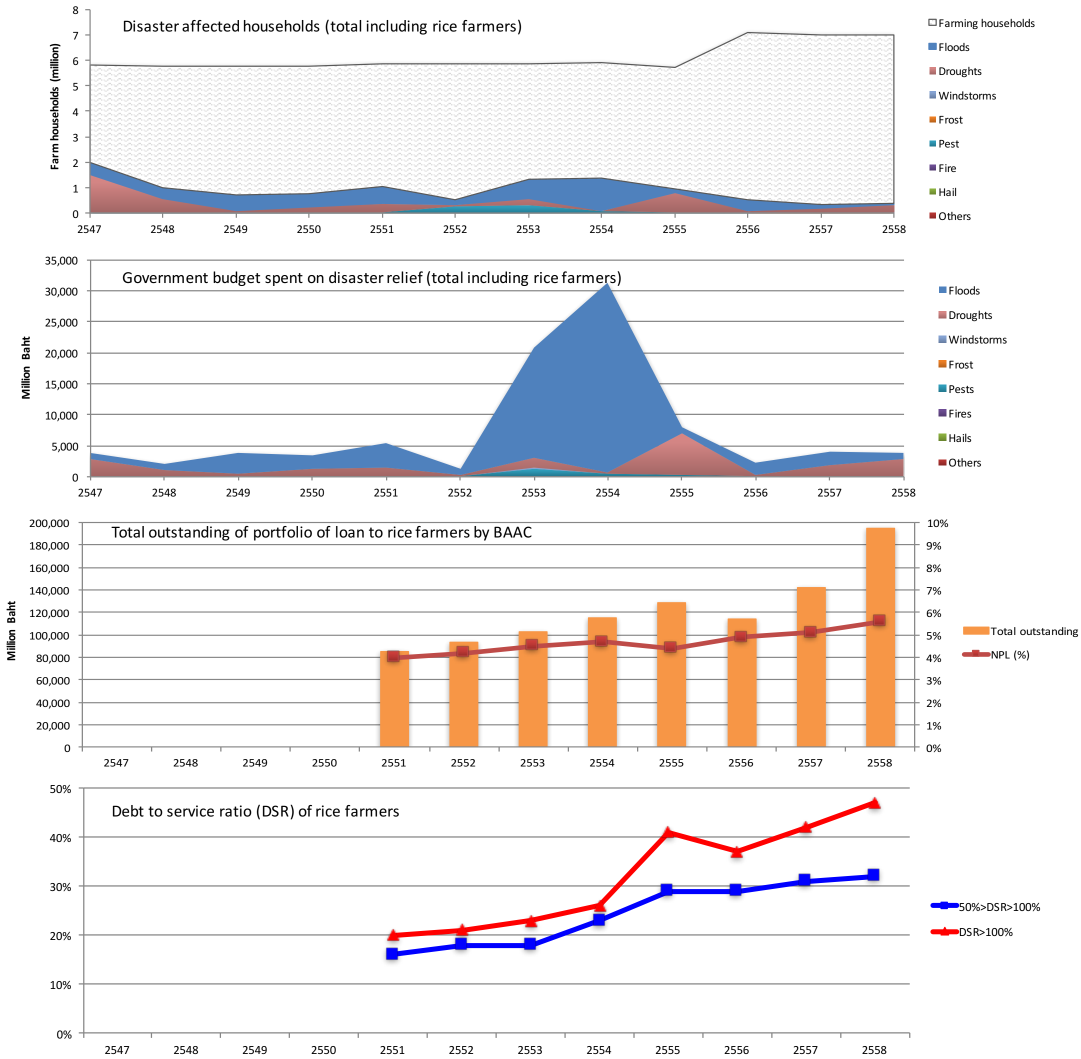
Historical statistics of rice areas affected by various types of disasters have been collected by the Department of Agriculture and Agricultural Extension (DOAE), who are also responsible for delivering disaster relief from the government. The data records total areas, affected farmers, number of affected tambons and total disaster budget spent by province. With the new data management system starting from 2554, it is also possible to digitalize and extract these loss statistics by Amphoe, though the data availability at that level are still largely inconsistent. Figure 3 thus summarizes loss statistics by province given 10-year historical data since 2548.

Figure 3: 10-year statistics of disaster losses by key risks



When disaster occurs, local government quickly check for large farm losses at their areas and report to provincial government for an announcement of disaster affected areas by the National Disaster Relief Program in accordance with the regulation of Ministry of Finance. Farmers in the affected areas declared disaster area are eligible to submit loss form to request for disaster relief. The loss is then verified by local government for requesting of relief. But since disaster occurs, it takes at least 30 days for loss verification to finish, another 15 days to report to Amphoe government and provincial government. If government budget is available at provincial level, relief disbursement can be done with in 15 days, making it possible to receive relief in 60 days after disaster. But the time usually is longer especially for the large disasters which would need budget from central government. Overall, a survey by Ministry of Finance reviewed that farmers felt that the relief suffer great delay (in some cases up to a year) and inadequate with relief payment of 1,113 baht per rai of up to 30 rais of affected areas covering about 30% of input cost.

Figure 4 Affected farmers, relief budget, loan portfolio outstanding and farmers' ability to repay debt



Loan portfolio of rice farmers by branch of the Bank of Agriculture and Agricultural Cooperatives (BAAC) from 2551-present are also obtained to understand how shocks/disasters could affect financial stability of the bank. Household production, losses and economic condition of households are also obtained from Agricultural Household Survey from the Office of Agricultural Economics covering about 3,000 rice farmers nationwide per year for 11 years from the production year 2547-48 to 2557-58. The data are repeated cross sectional at farmer level but with many villages being surveyed more than one time making it possible to construct (unbalanced) village pseudo panel data.

At the first glance, Figure 4 contrasts more macro evidence of total number of farmers affected by disaster, the corresponding government budget along with the growing loan portfolio outstanding of BAAC and capacity of farming households in repaying debt, proxy by debt to service ratio (total debt/total income) per year. This evidence review that despite the large fluctuation of government budget for relief, which could increase to more than 30,000 million baht in key disaster years, farm households still lack effective risk management and coping tool, resulting in welfare losses and potentially decreasing ability to relay their seasonal loan, which could in turn reduce financial stability of BAAC. The total loan outstanding that grows over time at an increasing rate reviews the potential of increasing loan default, delayed payment and government policy in debt holiday/restructuring for disaster affected farmers.

Current inefficiencies in agricultural insurance

As of now, there is one agricultural insurance program for rice farmers. The nationwide rice insurance which government subsidizes part of its premium is the micro-level product. This so called "Rice Disaster Relief Top-Up Insurance Scheme" (hereinafter Scheme) is the indemnity-based product. Two criteria must be met before a claim can be paid. Firstly, the farm must lie within an area declared as a disaster area by the local governor. Secondly, the entire farm or a proportion of it must be declared a total loss under the loss assessment procedures for the National Disaster Relief Program in accordance with the regulation of Ministry of Finance. Bank of Agriculture and Agricultural Cooperatives (BAAC) has been the only distribution channel in this scheme.

The Scheme was first launched to farmers in 2011. Table 1A and 2A in Appendix illustrate overview and historical performance of the Scheme respectively. In 2011, the single premium rate was 120 baht for the maximum coverage of 1,400 baht for flood, drought, windstorm, frost, hail, and fire perils. There were 8 private insurers participating the scheme. The second year of the insurance scheme, the premium rate in 2012 was maintained at 120 baht per rai, and the coverage was 1,111 baht per rai for the previous 6 natural perils, and 555 baht for pests and diseases. The only insurer for the 2012 scheme was the National Catastrophe Insurance Fund (NCIF).

While the Scheme was still in its infancy insuring less than 2% of the main rice area and experienced anti-selection together with severe flood in 2011 and drought in 2012, the Scheme suffered high loss ratios over 500% in 2011 and 2012. Consequently, risk rating for five zones were introduced. From 2013 to 2015, different premiums were applied for each zone, depending on their historical damage ratio at provincial level under the National Disaster Relief Program (Table 3A in Appendix). The premium rates for five zones ranged from 11% to 43% for the coverage of 1,111 baht per rai. With this risk rating, the private insurers and international reinsurance market regained confidence to take up the portfolio. In 2016, the government decided to bundle the rice insurance produce to the BAAC borrowers which totals 25 to 30 million rais. As a result of this huge increase in numbers and better spread of risk, a single risk premium of 100 baht per rai will be charged for all 2016 policies.

In all years except 2016, the premium was partially subsidized by government (50-75% and increasing in risky areas with higher premium) resulting in great reduction in premium, which still however could not attract enough demand for the product. In 2016, government and BAAC jointly pay for the full premium to all the farmers. Would this support suitable and would this create reverse incentive to farmers to protect themselves?

Simplicity of the Scheme is the key strength that government decided to deploy. However, there are rooms for improvement for sustainability. There are however many challenges that create inefficiencies in the current program. First, the scheme lacks high quality risk information of farmers with long historical profiles. As mentioned, the Scheme solely relies on the data recorded by loss adjustors who performed the loss assessment procedures for the National Disaster Relief Program in accordance with the regulation of Ministry of Finance. For the premium ratemaking process, the longest data series on historical loss reports recorded by Department of Agriculture and Agricultural Extension (DOAE) are up to 2005 and are limited only at the provincial level. This creates obstacle to accuracy in risk-based pricing of the insurers and potentially face adverse selection problem when the program tends to attract risky farmers. And so the insurer (without enough information) would have no choice but to charge high premium to protect themselves from potential losses, hence making premium rate currently very high.

Second, current loss verification process is less transparent (at least to the insurers) and suffer great delay and potential basis risk problem (i.e. when insured farmers experiencing losses but are not eligible to claim indemnity payment). For the claim administration, insurers can only wait for the loss assessment reports recorded by DOAE. The report process is timely and cannot be monitored by insurers. For insurance policy design, since the Scheme ties with the National Disaster Relief Program, a farm where it does not lie within an area declared by the local governor as a disaster area will not be eligible for payouts. It creates basis risk to an insured who experienced loss in his farm but does not get any payouts, even though they have been informed clearly by BAAC about terms and conditions beforehand. Both the long delay in payout and potential basis risk decrease value of this insurance reducing demand for the product. The less transparent loss verification further push insurers to increase premium to protect themselves from 'unknown risk' they could not observe.

These inefficiencies initiated by the lack of high quality risk data and loss verification resulted earlier in less than 2% of farmers insured but increasing premium (and so burden to government in subsidizing premium). The lack of effective market for insurance have made BAAC (who lends regularly to farmers) the place for farmers to transfer risk (through late loan repayment or default) resulting in low quality of loan portfolio and financial performance.

How high-quality risk information that can accurately, transparently measure losses of farmers with rich historical records and at low cost and in near rea time can unlock these inefficiencies? Such data can help insurers observe actual risk of insurable farmers, verify losses and eligibility of insured farmers in transparent manner, allowing them to do risk-based pricing accurately without having to load premium rates for 'unknown' factors and perhaps increase their incentive to design more variety of products to offer not just farmers but groups of farmers or BAAC. This perhaps can increase competition to insurers, which are the key to development of healthy insurance market. More accurate and near real time loss verification also potentially increase value of the insurance to farmers, potentially increase demand for the product and at the same time their ability to manage risk. BAAC and government can also use these data to target assistant better to ensure that the help would not destroy incentives of farmers in taking care of themselves.

The ecosystem of current risk management of farmers with limited information and so less-developed insurance market and the potential value of better risk information are summarized in Figure 1A in Appendix.

2. Recent development of satellite technology for agriculture

Satellite technology or remote sensing can be described as activities of recording/observing/perceiving objects of events at distance places. Therefore, the sensor receives this information without in direct contact with observed objects or events. The information is generally carried to the remote sensor by electromagnetic radiation via the atmospheric mediums. The output of the system is usually an image collecting observed information. Then, image will be brought to extract the useful information by interpreting and transform it to spatial data. This technology has been developed since World War II in due to military purpose. After that, it is gradually evolved into a scientific subject and released to public. The development of remote sensing platform and sensor has continued since then until nowadays with an initiation of airborne remote sensing developed to satellite platform. The resolution, both spatial and radiometric, are getting more quality and have various systems to suit the thematic purposes and many series have long historical profiles. At the present, we can access to remote sensed data independently and there are many distributors providing services both commercial and for free with different resolutions.

Remote sensing has been widely used in many field, especially in agricultural activity. This technology is adopted by many organizations, both government and private companies. In Thailand, rice is an important cash crop which mainly plant in all region. Many remote sensing techniques are adapted to rice field for examples, classification, change detection, crop monitoring, etc. to support and promote rice planting, both quantity and quality. Multidiscipline algorithms related to rice field are continue developed to improve the ability of remote sensing on rice study.

According to electromagnetic radiation which is playing as an information carrier, there are two main sensors related to acquiring ground information. The first one is Passive remote sensing. This system uses the sun radiation as an information carrier. Electromagnetic wave from the Sun travels through the atmospheric window to objects on the Earth and the energy is reflected and emitted. After that, energy returns to the sensor on the space platform as signal and then, transfers information to ground stations for digitally collected. Products from passive remote sensing will be described as optical image which is contained visible to infrared wavelength (multispectral band). Image could be collected both wide and narrow wavelength due to the sensor. Generally, wide wavelength will be collected into three visible bands (red, green, and blue) plus infrared bands. In case of narrow wavelength, it is called hyperspectral and could have more than 100 bands collected as a cube. The spatial resolution is varied from half a meter to many kilometers due to the sensors' design which is related to the size of area coverage. However, passive remote sensing requires sun radiation. Then, this sensor could collect ground information only in daytime. Moreover, there are many obstructions in atmosphere for examples, cloud, aerosols, water vapor, etc. which can reduce the ability of image acquisition to this platform.

The other sensor is active remote sensing which can produce the signal itself transmitting to the Earth's surface and returns the scattering signal to the platform. This sensor is generally called Synthetic Aperture Radar (SAR) remote sensing. Due to it can create high frequency wavelength or Microwave itself, SAR has ability to perform image collection both day and night. Microwave can travel through cloud and any particle matters in the atmosphere which is regardless to cloud cover. However, the analyzed method for this sensor is quite complicated.

Satellite image provides many advantages to rice field study. The multi-spectral capability helps us to obtain a near infrared wavelength which is best reflect the cellular structure of vegetation and can be derived the Normalize Difference Vegetation Index (NDVI). The index has can investigate cropping areas and crop's health at various stages of crop. The index also comes with more than 16 years of continuous historical profile. Due to its revisit capability, multi-temporal has been applied to rice field study, especially on extracting rice phenology, to measure the characteristic of rice growing stage. This is very useful to monitor crop planting and analyze how the extent of areas appearing with rice at different growing stage. In case of disaster event, satellite can be used to indicate losses to ensure timely humanitarian responses.

However, as mentioned above, optical image also has some energy scattering and almost occur in the troposphere layer. Moreover, cloud also has much effect on acquiring image, especially in tropical region where cloud covers area most of the year. It can be corrected using the atmospheric correction algorithms. Alternatively, in tropical region, SAR is selected instead of optical sensor to overcome cloud effect, though one needs to tradeoff with the shorter historical profile.

Remote sensing applications has been widely used in agriculture. Some notable example includes USDA-FAS. It produces the map of crop growing mainly in the United States and all over important growing regions. Information provided on USDA web service has a lot more precision. This is a result of systematically processes and many months of survey. USDA's officers are assigned to collect data in each county for example, crop type, planting date, acreage, crop's yield, pest invasion, etc. and submit this information to analyze with satellite's data. NDVI information is extracted every revisit period to construct the vegetation map and meteorology data is acquired from weather stations all over US. After combining data (physical and remote sensing) together, each crop type's area will be classified and estimate planting area, crop stage, and crop yield which will impact the CBOT market. Moreover, crop monitoring is operated by these data collections. USDA will analyze the probability of disaster that will impact on crop and estimate the loss. This information could then be used to design various risk management instruments including crop insurance, where insurers can use satellite imagery to verify eligibility of the insured, monitor farming activities and verify farm losses for fast indemnity settlement.

The other good practice is operated by French private company named "GEOSYS". They create many systems related to remote sensing data for examples Agriquest and Farmsat. Agriquest is a crop monitoring system which provides NDVI, precipitation, temperature, solar radiation, growing degree day information with crop expert's analysis. Furthermore, the system can alert the warning when it detects some abnormal situation. Farmsat is precision agricultural system to support farmers to growth their crop and manage their planting. The system manages all activities from preparing bared land until harvesting. Farmers will receive information and warning via interface and can prepare if there is any problem occurs in their farm. Farmers can control their budget on factors of production and estimate their income at the end of the growing season. Moreover, they can use modern technology to support farming which can reduce cost and time consumption. It would have more efficiency if farmers can set up the cooperation to contribute the benefit of technology to each other. Advanced technology makes farming easier than ever.

In Thailand, there are also many initiatives that combine remote sensing and GIS information for agriculture and disaster risk management. Various initiatives at GISTDA include *flood monitoring*, which use various remote sensing products (Radarsat and various optical images including Thai Earth Observation System (THEOS), QuickBird, WorldView and Landsat data) and provides information on history of floods as well as the on-going and expected flood events. *G-Agro* is another platform developed to provide near real time updates and monitoring of rice production for estimation of agricultural yields and market management. Similar initiatives also expand into key economic crops of Thailand. *GISAGRO* is another very useful platform that facilitate registration of farmers, allowing them to fill in GIS locator of their field, to receive valuable planting, input information necessary for making planting decision. Another powerful initiative in Thailand includes *Agri-map*, developed in cooperation between the Ministry of Agriculture and NECTEC for dynamic management of farmland. The Agri-Map for managing agriculture in each province was made in accordance with the current and future situation concerning production, supply and demands of farm goods. The program will also demonstrate whether the use of farmland of each province is appropriate to the type of soil, water volume and types of plants in order to make adjustments and balance so that the farmers' problems will be correctly solved. Some detail of these initiatives can be found in Appendix 1.

3. Can satellite data be used to generate better risk information for Thai rice farmers?

Satellite data and processing

We utilize two remote sensing data available at GISTDA: (1) the Terra MODIS NDVI at 250m resolution available every 8 days from 2001-2016 and 2) the RADARSAT at 50m resolution available from 2011.

The moderate resolution imaging spectroradiometer (MODIS) is the polar orbiting satellites that have been comprehensively used such as, classification land cover from MODIS EVI (Xia et al., 2008), detecting land cover change from MODIS NDVI (Kleynhans et al., 2011), using MODIS to map sea surface oil (Innman et al., 2010) and drought monitoring from MODIS LST (Fan et al., 2012). The normalize difference vegetation index (NDVI) and the enhanced vegetation index (EVI) are the standard vegetation products from the MODIS. These indices are crucial for effectively characterization of bio-physical/ biochemical states and processes from vegetated surfaces (Silveira et al., 2007). Many applications such as remote sensing image classification, land cover change detection and monitoring crop phenology rely on the information derived from these indices. In this research, the NDVI product is chosen to be used in the estimation of the cultivation date.

The main reason that we choose the NDVI rather than EVI is due to fact that the NDVI data have a resolution of 250 meters (approximately 62.5 rai per pixel) whereas the original EVI data have only resolution of 500 meters (approximately 125 rai per pixel). The area of 500*500 meters is too coarse for the effective rice monitoring system in Thailand.

The time-series of NDVI data has been popularly used in continuously monitoring vegetation phenology which is seasonal plant cycle stages and land cover characteristics. Although the NDVI time-series cannot discern specific phonological stages, such as bud burst or flowering, they can successful describe the seasonal dynamics of vegetation. Nevertheless, the MODIS date are suffered from various types of noise and degradations due to the atmospheric conditions. In particular, the MODIS satellite cannot observe the earth surface when high cloud cover is present in the scene. Therefore, noise reduction is necessary before data are used. They have many approaches to address this problem such as applying the savitzky-golay filter, fourier transform, wavelet transform, markov model, Extended Kalman filter to the time-series NVDI data to smooth out the atmospheric effects.

Additionally, there are many studies that employ MODIS data to estimate phenology date. For example, first Sakamoto et al., 2005 proposed a crop phenology detection method using EVI every 8 day at 500m resolution. Crop phenology dates include planting, heading and harvesting dates; heading date can detect by the maximum EVI in the time profile; planting date can detect by decreasing heading date 60 days. And first and second derivatives equal 0; harvesting date can detect by increasing heading date 30 days. And second derivative equals 0. Second, (Hu et al., 2009) proposed crop phenology date estimation method from NDVI daily data derived from MODIS 250m resolution to detect emerge, silking, dough, maturity and harvest dates in corn and emerge, bloom, setting pods, drop leaves and harvest dates in soybean. To estimate phenology date, first and second derivatives of smoothing NDVI time-series data can extract dates by finding the silking date of corn and bloom date of soybean, which NDVI reach to the high value range and first derivative is near zero, and then finding the others dates by relation of these date and derivatives of data approximately to zero. Last, (Tan et al., 2011) proposed vegetation phenology estimation method by using TIMESAT algorithm to smooth MODIS EVI data every 8 days at 500 m resolution and then using first, second and third derivatives of plantation period of EVI time-series data to extract date. Summary, their method cannot detect planting date as soon as possible. They have to find the date at the maximum NDVI first. Thus, our point is the estimating the cultivation date without knowing the maximum NDVI date.

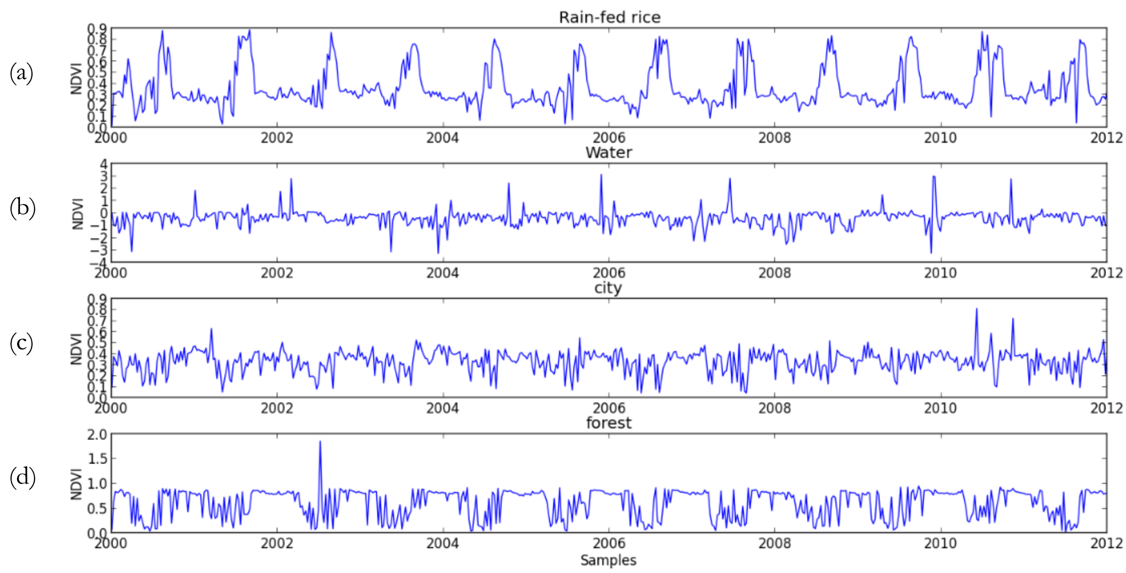
The 8-day MODIS/Terra data (MOD09Q1) spanning from 2001 to 2016 were freely downloaded though the Earth Observing System Data Gateway. It has two spectral bands for red and near-infrared spectral bands with a spatial resolution of 250 meters. The data were formatted using a Sinusoidal projection, but were re-projected to the Geographic coordinate, mosaicked, and subset over Thailand.

The NDVI is used to study the vegetation growth estimation is normalized transform of the near-infrared to red reflectance ratio follow as

$$NDVI = \frac{(NIR - RED)}{NIR + RED} \tag{1}$$

where *NIR* and *RED* are the digital numbers in near-infrared and red spectral bands, respectively. Since vegetation reflect the NIR color spectral better than the red color, the NDVI has a higher value if there is more vegetation, and NDVI is low when a scene does not have any vegetation. Usually, the NDVI data is between -1 and 1 where 1 and -1 indicates full vegetation and no vegetation, respectively. To create the time-series dataset for this 16 year period; the NDVIs for every 8day MODIS scene were computed from the 8-day MODIS data. There are in total of 747 NDVI data points for each pixel from June 2001 to July 2016. The examples of time-series NDVI data are show in Figure 5.

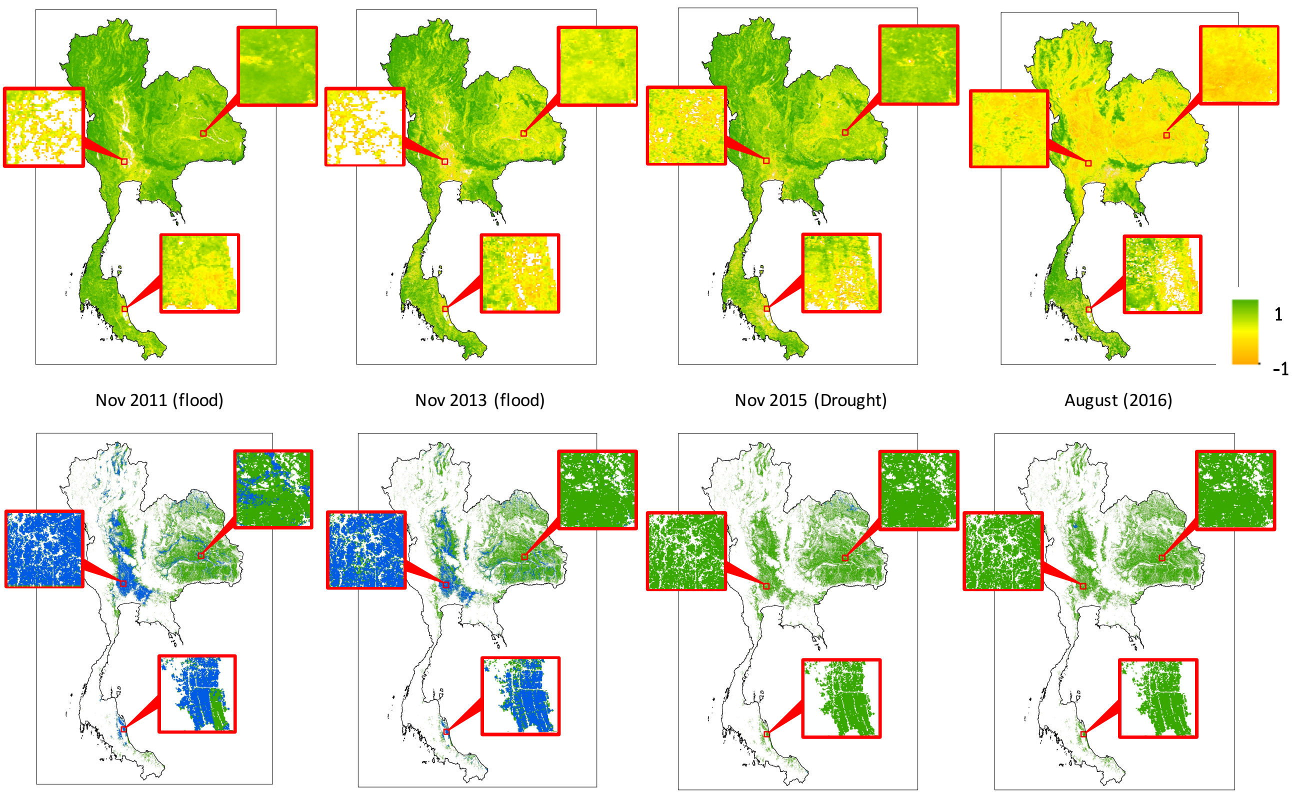
Figure 5: NDVI pattern; (a) Rain-fed rice; (b) Water;(c) City; and (d) Forrest.



Radarsat data set spanning from September to December 2011 was used in this flood monitoring effort. Both Radarsat-1 and Radarsat-2 satellites are equipped with C-band SAR sensors (5.3 GHz for Radarsat-1 and 5.4 GHz for Radarsat-2). Radarsat-1 transmits horizontal polarization (H) and receives H data only (i.e. in combination HH data). Radarsat-2 can transmit both H and V polarizations and receive either H or V signals or both simultaneously. Radarsat antennas were designed to operate with multiple beam modes, giving fine to coarse resolution and variable inclination angles. This data will be combined with MODIS-NDVI data to better reflect the loss of rice areas by flood.

Figure 6 depicts variation of NDVI and Radarsat flood data for key disaster years and zoom in to show variation within the 10 kilometer squared rice areas. When crop experiences losses e.g., from flood or droughts, NDVI appears with very low value.

Figure 6: MODIS NDVI and RADARSAT flood data for present and disaster years



All 16 years of MODIS NDVI data and RADARSAT flood data are then extracted for all potential rice growing areas in Thailand identified earlier by GISTDA taking into account both satellite imageries and ground verification.

Detecting growing area and dates using MODIS-NDVI

Existing literature: Satellite images can be very useful in agricultural study, since a satellite image can cover the large scale area making it is easier to monitoring the agriculture in the large scale. One of the comprehensively use satellite images are Synthetic Aperture Radar (SAR) providing a global data set every 24 days, we collected the data set from RADARSAT-2 satellite. The spatial resolution of SAR depends on type of beam modes, we collected the data set on scanSAR narrow, there is a spatial resolution is 50 meter.

Suwannachatkul et al. proposed to use the MODIS data for the cultivation date estimation of the rice fields. In their work, the Savitzky-Golay filter is first employed to smooth the noisy NDVI data. Then, the rice phenology is categorized into four stages: pre-cultivation, growing, maturation, and harvest. They estimated the transition probabilities from the actual observed NDVI by using the Hidden Markov model, and created a trellis diagram to find the possible paths of the rice growth in time-series data with the Viterbi algorithm. The weather conditions in the rainy season reduced the accuracy of cultivation harvest date estimation (Suwannachatkul et al., 2014)

Similar work by Chumkesornkulkit et al. used the MODIS NDVI data for the cultivation date estimation. In their work, only rain-fed rice was considered. The main algorithm was based on the extend Kalman filter (EKF) to smooth the observed data by using the modulated cosine function consisting of three parameters, namely the mean, amplitude, and phase. The cultivation dates were estimated as the dates where the seasonal variation derived from the EKF was greater than a threshold after its minimum. The NDVI data in Thailand can be affected by clouds mostly in the rainy season, which is the cultivation period of the rice, but in other periods the effect of clouds is a lot less. This problem may influence the accuracy in cultivation date estimation. (Chumkesornkulkit et al., 2013)

For the work in SAR, Lopez-Sanchez et al. used the SAR Polarimetry at X-band to find the rice phenology by referring to Biologische Bundesanstatl, Bundessortenamt and Chemische Industrie (BBCH) stages. These BBCH can be categorized into three large stages, namely, vegetative, reproductive and ripening, in order to assign the BBCH to the majority of pixels inside the parcel, when the plant emergence (BBCH 18-21) and last stage (BBCH 50+) are still present in some areas. The Terra SAR-X has high spatial resolution but low coverage, because the swath width is around 15 km on the grounds and some measurements are below or close to the noise level of TerraSAR-X(-19dB). (Lopez-Sanchez et al., 2012)

Three RADARSAT-2 quad-polarization images from transplanting to rice crop harvesting stages were used by Wu et al. The authors also compared these images with the backscattering values and with the ground data such as rice height and dry biomass. Their research showed the relation of backscattering coefficient to plant height and backscattering coefficient to dry biomass. The satellite images with HV polarization had the highest correlation to the plant height and dry biomass. The ratio of image HH/VV can be used to separate the rice crop from banana plantations, forests, and rivers. (Wu et al., 2012)

Ribbes et al. showed that the backscattering coefficient of the rice field appeared to have a significant temporal variation and the highest correlation with age (day after sowing), plant height and plant biomass of rice in term $R^2 = 0.90, 0.81, \text{and } 0.80$, respectively. The trend of backscattering coefficient at the beginning cycle revealed that flooded fields provided low backscattering (-14 to -12dB), with the value reaching -6 dB until the end of the cycle. The vertical structure of the rice plants that attenuated the radar signal by the canopy was higher for VV than for HH polarizations. The accuracy of mapping was 87% compared with the available ground data. (Ribbes et al., 1999)

These various study illuminate the potentials of remote sensing in detecting rice area, crop cycle. Our work continues to further explore the extent to which remote sensing data can be used to detect farm losses of farmers.

This study uses a triply modulated cosine a factor function in the nonlinear estimation method that Kleynhans et al., 2010 has proposed. A triply modulated cosine function has equation as follows

$$y_k = \mu_k + \alpha_k cos\ (\omega k + \varphi_k) + n_k \tag{2}$$

where y_k is the observed value of the NDVI time-series at time k and n_k is the noise sample at time *k*. Here, the cosine function at the time k is modeled to have the angular frequency of ω, the mean value of μ_k, the amplitude of α_k, and the phase of φ_k. The angular frequency can be explicitly computed as $\omega = 2\pi f$, where f is based on the annual vegetation growth cycle in the area of interest. For example, in the rain-fed rice, we use $f = 8/365$ whereas, for the irrigated rice, $f = 8/240$ for 2- and 3-crop cycles per year.

The estimation of triply modulated cosine function parameters is nontrivial and requires a nonlinear estimator. A good candidate for parameter tracking for the non-linear system is the extended Kalman filter (EKF, following Kleynhans et al., 2010, 2011). The EKF has been used for detection and reconfiguration surface fault in aircraft control systems (Caliskan et al., 2000), improving accuracy of the feature map by modified Neural Network aided (Kang et al., 2010) and estimate the absolute position of an autonomous mobile robot (Marron et al., 2003).

According to the EKF formulation, for every increment of k (the discrete time), a state vector $\boldsymbol{x}_k$ is defined containing the parameters to be estimated in the form $\boldsymbol{x}_k = [\mu_k \quad \alpha_k \quad \varphi_k]^T$. The relation between $\boldsymbol{x}_k$ and $\boldsymbol{x}_{k-1}$ is given by a function $f(.)$ which is a known and can be a nonlinear function. The state vector $\boldsymbol{x}_k$ is related to the observation vector y_k via a nonlinear measurement function $h(.)$. Due to noisy nature of the measurement, both systems are corrupted with unknown noises, and their relationships can be written as

$$\boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_k \tag{3}$$

$$y_k = h(\boldsymbol{x}_k) + \boldsymbol{v}_k \tag{4}$$

for the estimations of the state and observation vectors, respectively. The terms w_k and v_k are process and observation noises, respectively. The state vector parameters may be estimated over time k by recursive iteration based on the observation data y_k up to the time k. In the observation equation (4), y_k is the predicted measurement.

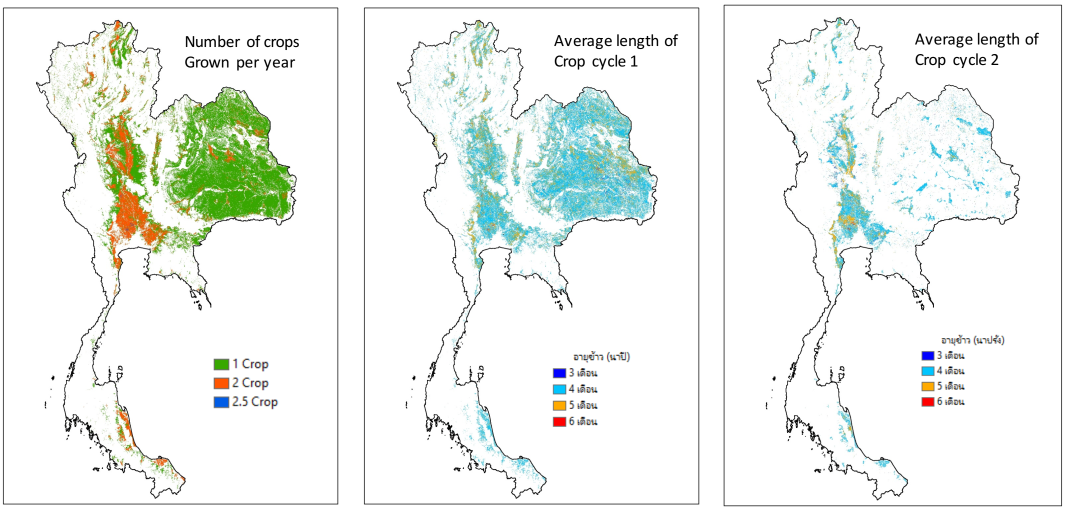
Interpretation of the term μ_k and $\alpha_k\, cos\,(\omega k + \varphi_k)$ of a triply modulated cosine function can be given that μ_k can be considered as the trend variation and should be associated with long term vegetation changes such as the deforestation whereas the term $\alpha_k\, cos\,(\omega k + \varphi_k)$ can be considered as the seasonal variation, and should be related to the growth of rice. As a result, we use only the term $\alpha_k\, cos\,(\omega k + \varphi_k)$ in the estimation of cultivation time in this study.

Before cultivation, the rice field is clear of vegetation as a result, the NDVI is low. As rice grows, more vegetation covers the rice field and the NDVI increases. However, in certain area, farmer put down some other crops such as green beans or grasses before cultivate rice to increase the Nitrogen in the soil. Hence, we identify the rice cultivation date as the first time instance that seasonal term $\alpha_k \, cos\,(\omega k + \varphi_k)$ is greater than a predefined threshold τ after the lowest point.

We employed this technique to detect planting date in each season of each and every pixel of NDVI data throughout the country. Figure 7 depicts average number of crops grown per year, length of each crop season for the rice growing areas in Thailand

This exercise thus allows us to detect planting dates for each cropping season each year and so can estimate the corresponding planting areas. Next, we will take this result, explore variations in NDVI and cropping patterns to identify homogenous rice production zones of the country so the loss estimation based on MODIS-NDVI and RADARSAT can be done similarly within each zone but could differ across zones.

Figure 7 Detected pattern of crops and rice crop cycles



Clustering homogenous rice production zones

K-mean cluster analysis is then used to separate rice production areas into 15 homogenous zones using GIS information, rainfall information and NDVI data and so based on geographical characteristics including irrigation, altitude and rice species, meteorological information including mean and standard deviation of rainfall, and key characteristics of NDVI in each pixel including mean value, standard deviation, curvature, number of detected crops grown per year, detected length of first crop, detected length of second crop. Results are reported in Table 1.

Table 1 Summary statistics of 15 homogenous rice production zones

Cluster	Cluster name	NDVI		Number of	Crop 1 cycle	Crop 2 cycle	Irrigation = 1	Rice species	Elevation
		Mean	SD	crop per year	(months)	(months)			(meters)
1	Central - rainfed	0.53	0.12	1.00	4.18	-	0.15	Mixed	39.77
2	Central - irrigated - jasmin rice	0.57	0.13	2.01	4.41	3.20	0.85	Jasmin	33.85
3	Central - irrigated - low variations	0.59	0.12	2.01	4.27	3.18	0.97	Mixed	9.56
4	Central - irrigated - high variations	0.55	0.18	2.00	4.41	3.51	0.96	Mixed	9.38
5	North - rainfed	0.53	0.14	1.00	4.24	-	0.08	Mixed	170.15
6	North - irrigated - jasmin rice	0.58	0.13	2.01	4.20	3.19	0.85	Jasmin	324.24
7	North - irrigated - low variations	0.58	0.12	2.02	4.47	3.09	0.70	Mixed	77.54
8	North - irrigated - high variations	0.52	0.19	2.00	4.45	3.35	0.74	Mixed	74.27
9	NE - rainfed - jasmin rice	0.48	0.13	1.00	4.66	-	0.05	Jasmin	172.14
10	NE - rainfed - sticky rice and others	0.50	0.12	1.00	4.73	-	0.06	Sticky	210.17
11	NE - irrigated jasmin rice	0.51	0.14	2.00	4.26	3.04	0.79	Jasmin	170.05
12	NE - irrigated others	0.54	0.13	2.00	4.46	3.06	0.60	Mixed	208.77
13	South - rainfed	0.66	0.10	1.00	4.02	-	0.06	Mixed	25.39
14	South - irrigated - low variations	0.65	0.10	2.00	4.04	3.03	0.98	Mixed	13.46
15	South - irrigated - high variations	0.57	0.13	2.00	4.07	3.27	0.93	Mixed	4.95

Figure 8: Homogenous rice production zones identified by cluster analysis

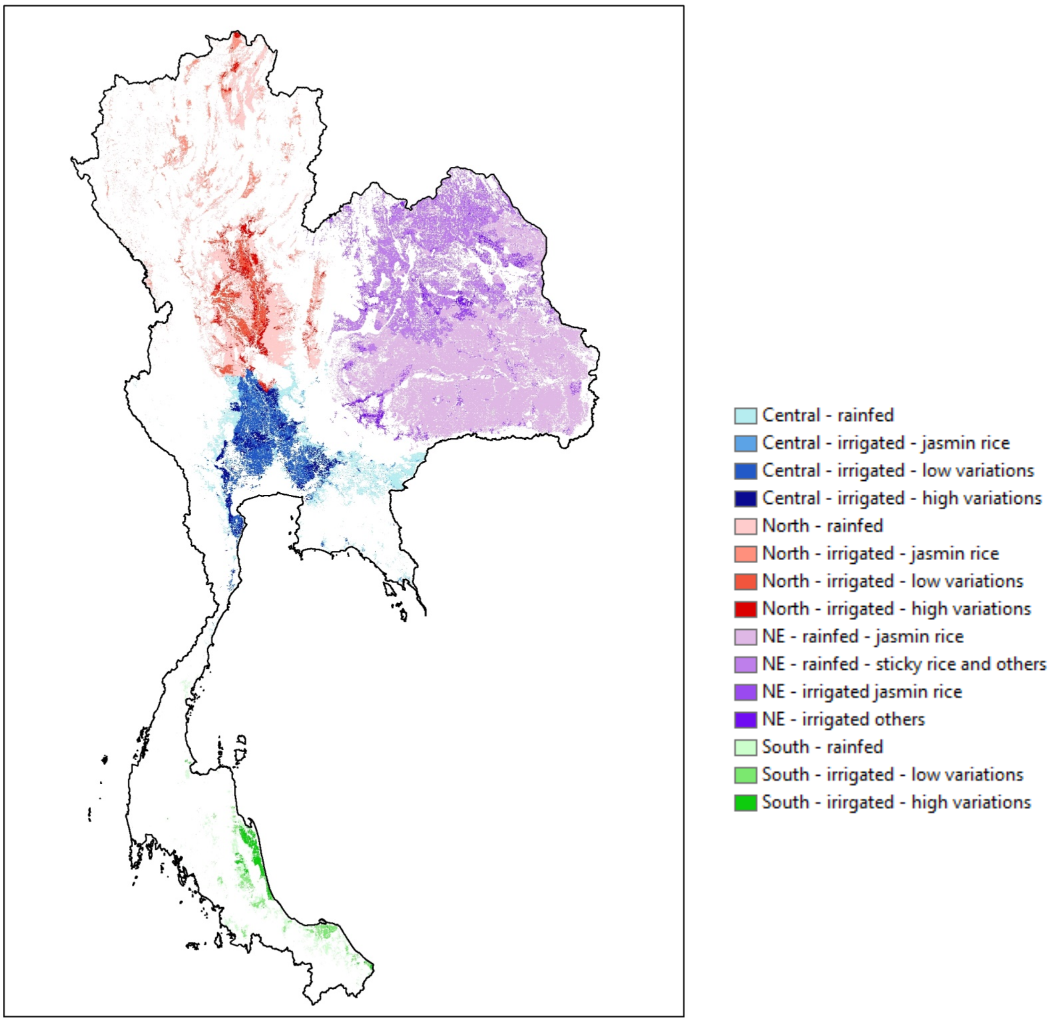
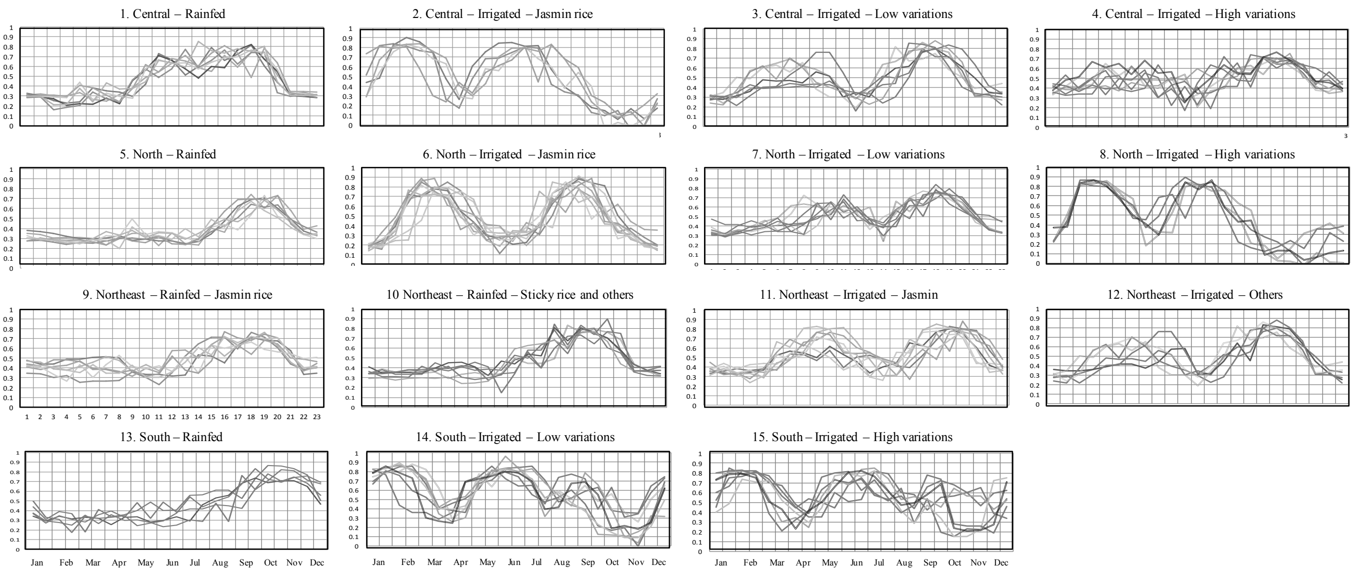


Figure 9: Distinct patterns of NDVI across 15 rice production zones



These 15 homogenous production zones with distinct NDVI patterns form the basis of further analysis on farm losses.

Estimating production losses

We estimate production losses by detecting deviation of NDVI patterns from 'normal' condition specific to each pixel and production zone. In particular, NDVI should reflect some varying forms of bell shape when farmer starts planting. We then simply divide growth cycle from the detected planting date into 4 stages each with 1 month, 2 months, 3 months and 4 months following the planting date, making the full crop cycle of about 4 months or 120 days. In some production zones, especially in the rainfed areas of the Northeast, crop cycle could last up to 5 months and so additional stage is considered for those zones.

Production loss is defined as a composite function of deviation from 'normal' condition in each planting stage since the planting date is detected. Due to different level and nature of NDVI across pixels and production zones, definition of 'normal' should be pixel and production zone specific. We thus define 'normal' level by pegging the level of NDVI observed in each crop stage as some percentile value based on historical data observed in that pixel. The level of percentile will then vary across production zones but similar for all pixels in the same zone. Since cloud effect of the 8-day data (resulting in constant drop of NDVI to 0) could present key threat to this exercise, we rather convert the 8-day composite to 16-day composite taking the maximum values of the 8-day data.

In particular, production loss in pixel p of crop c in year y can be calculated based on the following combination of 16-day NDVI in any continuous 2-week interval t as

$$\tilde{L}^{c}_{py}(n^{*}_{pt}, n^{**}_{pt}, w_s) = \begin{cases} \displaystyle\sum_{plant(ndvi_{py})<c<harvest(ndvi_{py})} w_s E_{t\in s}\left(\frac{n^{*}_{pt} - ndvi_{pty}}{n^{*}_{pt}}\right)\mathbb{1}(ndvi_{pty} < n_{pt}) \\ 100\% \;\; if\ in\ any\ growing\ stage\ s,\;\; E_{t\in s}\left(\frac{n^{*}_{pt} - ndvi_{pty}}{n^{*}_{pt} - n^{**}_{pt}}\right) = 100\% \end{cases} \tag{5}$$

where $plant(ndvi_{py})$ represents the detected planting date based on series of 16-day NDVI in that pixel p and year y and $harvest(ndvi_{py})$ represents the detected harvest date based on the same series in pixel p and year y. $E_{t\in s}\left(\frac{n^{*}_{pt}-ndvi_{pty}}{n^{*}_{pt}}\right)$ reflects mean value of deviation from 'normal' condition in each crop stage s which can be derived from the deviation of the two of 16-day NDVI data during that crop stage from some percentile threshold n^{*}_{pt} based on historical profile of NDVI in that pixel and that 16-day period t since planting. $\mathbb{1}(ndvi_{pty} < n_{pt})$ is an indicator function which will be equal to one if $ndvi_{pty} < n_{pt}$ and zero otherwise. The function thus allows us to only take into account the NDVI values when they dip below 'normal' condition. Overall, total production loss is thus some combination of below normal condition of each and every growing stage in each crop season based on some weighting scheme w_s that defines relative contribution of abnormal NDVI in each particular stage to total production loss in that season.

The production loss function allows for potential of total loss that could occurs before the season ends by capturing the extreme fall of NDVI from 'critical level' defined as some percentile of NDVI in each particular pixel and 16-day period since planting n^{**}_{pt}.

The estimating production loss $\tilde{L}^{c}_{py}(n^{*}_{pt}, n^{**}_{pt}, w_s)$ is thus a function of three parameters $n^{*}_{pt}, n^{**}_{pt}, w_s$ that defines 'normal' condition and 'critical' condition in each crop stage of each pixel and time, and the relative weight of each cropping stage toward total loss calculation. Figure 10 depicts framework of our loss calculation.

The production loss over an area a can then be calculated as

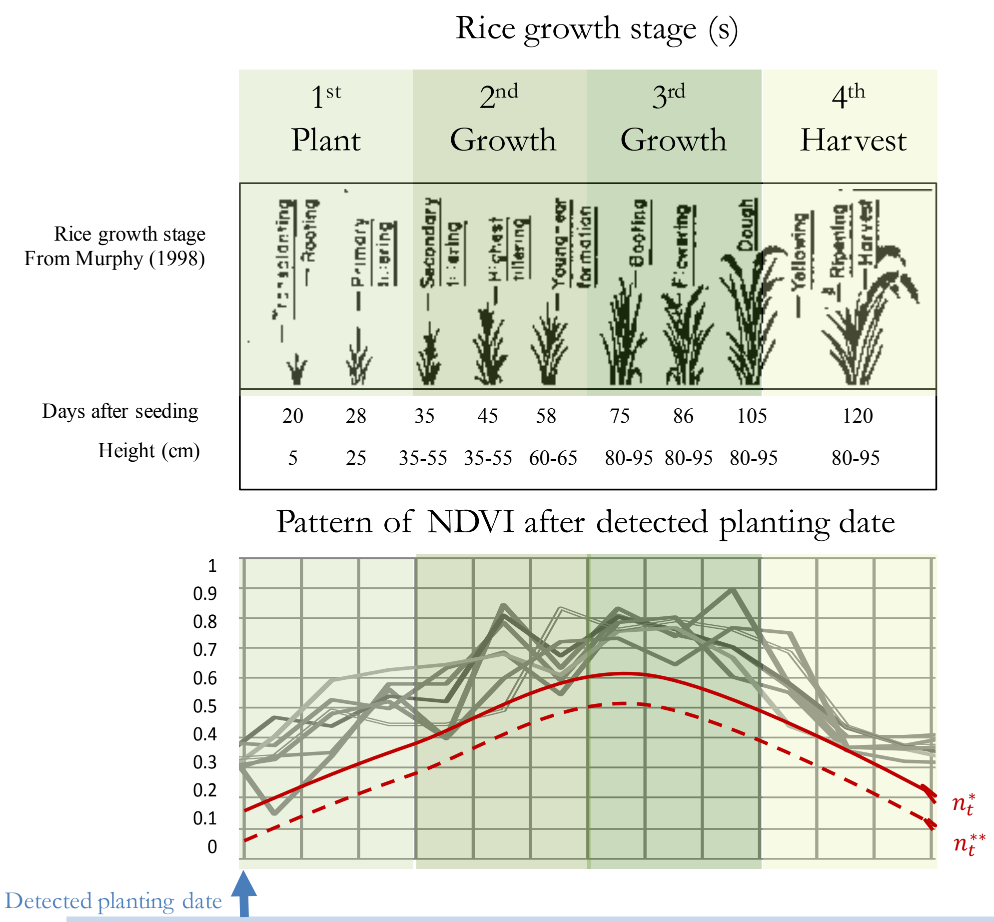
$$\tilde{L}^{c}_{ay}(n^{*}_{pt}, n^{**}_{pt}, w_s) = E_{p\epsilon a}(\tilde{L}^{c}_{py}(n^{*}_{pt}, n^{**}_{pt}, w_s)) \tag{6}$$

We define production loss of individual farmer i with farm locating in an area a where NDVI data can be taken as

$$l_{iy}^{c} = \tilde{L}_{ay}^{c}\left(n_{pt}^{*}, n_{pt}^{**}, w_{s}\right) + \varepsilon_{iy}^{c} \tag{7}$$

where the variance of ε_{iy}^{c} thus represent 'basis risk' defined over the deviation of farmer's actual loss from the estimated loss from NDVI data.

Figure 10 Detecting production loss based on NDVI patterns



$$\text{Loss} = \begin{cases} w_1 E_{t\in 1}\left(\frac{n_t^{*}-ndvi_t}{n_t^{*}}\right) + w_2 E_{t\in 2}\left(\frac{n_t^{*}-ndvi_t}{n_t^{*}}\right) + w_3 E_{t\in 3}\left(\frac{n_t^{*}-ndvi_t}{n_t^{*}}\right) + w_4 E_{t\in 4}\left(\frac{n_t^{*}-ndvi_t}{n_t^{*}}\right) & \text{if } ndvi_t < n_t \\ 100\% \; if \; in \; any \; growing \; stage \; s, \; E_{t\in s}\left(\frac{n_t^{*}-ndvi_t}{n_t^{*}-n_t^{**}}\right) = 100\% & \end{cases}$$

Using the above framework, we estimate the optimal loss function by choosing the three parameters for each production so as to minimize the basis risk and so choose the three parameters for each production zone such that

$$n_{pt}^{*}, n_{pt}^{**}, w_{s} = Arcmin_{n_{pt}^{*}, n_{pt}^{**}, w_{s}} E\left(l_{iy}^{c} - \tilde{L}_{ay}^{c}\left(n_{pt}^{*}, n_{pt}^{**}, w_{s}\right)\right)^{2} \tag{8}$$

where we first obtain individual farmer's actual production loss data from the repeated cross sectional agricultural household survey of around 2,845 rice farm households nationwide resided in 610 tambons and 712 villages per year for 11 years from the production year 2547-48 to 2557-58. So altogether, the estimation is based on 31,295 farm households from every province. Table 2 provides summary statistics of the household data used.

Table 2: Summary statistics of agricultural household data used in the study

Variable	Mean	Std. Dev.	Min	Max
Demography				
Household members	4.7	1.7	1.0	10.0
Age	37.2	10.7	1.0	81.5
Dependent ratio (%)	26%	22%	0%	100%
Highest edu - Less than high school (%)	67%	24%	0%	100%
- High school (%)	22%	22%	0%	100%
- Vocational (%)	6%	13%	0%	100%
- Bachelor or higher (%)	5%	12%	0%	100%
Agricultural production				
Operating farm land (rai)	27.2	25.1	0.0	423.6
Number of rice crop grown per year	1.3	0.5	0.0	2.0
Rice crop 1 planting area	22.3	22.6	0.0	331.5
Rice crop 2 planting area	9.4	20.0	0.0	331.5
Rice crop 1 yield (kg/rai)	457.7	283.0	0.0	10,000.0
Rice crop 2 yield (kg/rai)	640.3	389.5	0.0	10,000.0
Total annual net income (Baht)	104,246.0	775,347.4	-1,926,500.0	3,165,820.0
Share of rice income to total income (%)	45%	23%	14%	93%
Financial portfolio				
Total loan outstanding (Baht)	128,112.7	218,341.7	0.0	5,000,000.0
Share of loan in agriculture (%)	54%	29%	27%	100%
Loan collateral - individual (%)	36%	43%	0%	100%
Loan collateral - group (%)	41%	43%	0%	100%
Loan collaterial - asset (%)	20%	32%	0%	100%
Saving (Baht)	39,174.9	198,448.5	0.0	11,100,000.0
Other asset (Baht)	1,721,390.0	3,331,409.0	0.0	143,000,000.0

Note: Average number of households per year is 2845 from 610 tambons and 712 villages. The data covers 10 years from 47-48 crop production year to 57-58 crop production year.

The data come with tambon location of each farmer and so we then match each farmer with NDVI data extracted at tambon level. Search algorithm is employed to then find the combination of the three parameters $(n_{pt}^{*}, n_{pt}^{**}, w_s)$ that minimize the basis risk for each homogenous production zones. Table 3 thus reflects search results of optimal combination of the three parameters.

Using these estimated parameters, we then constructed MODIS-NDVI predicted losses for each and every pixel over time. RADARSAT data is then overlaid over all the pixels to detect which pixels, when in the crop cycle when we see inundation in RADARSAT (and for how many weeks do we see inundation). We then use this information to cross verify with the estimated loss from MODIS-NDVI. And in the case when there is no indication of loss from MODIS-NDVI in any pixel while we observed more than 2 weeks of inundation from RADARSAT during the cropping cycle, we re-assign loss to those pixels.

Overall all, these algorithms thus allow us to construct and track crop production cycles and losses for each and every pixel of rice area as depicted in Figure 11.

Figure 11 Detected crop cycles and loss based on NDVI

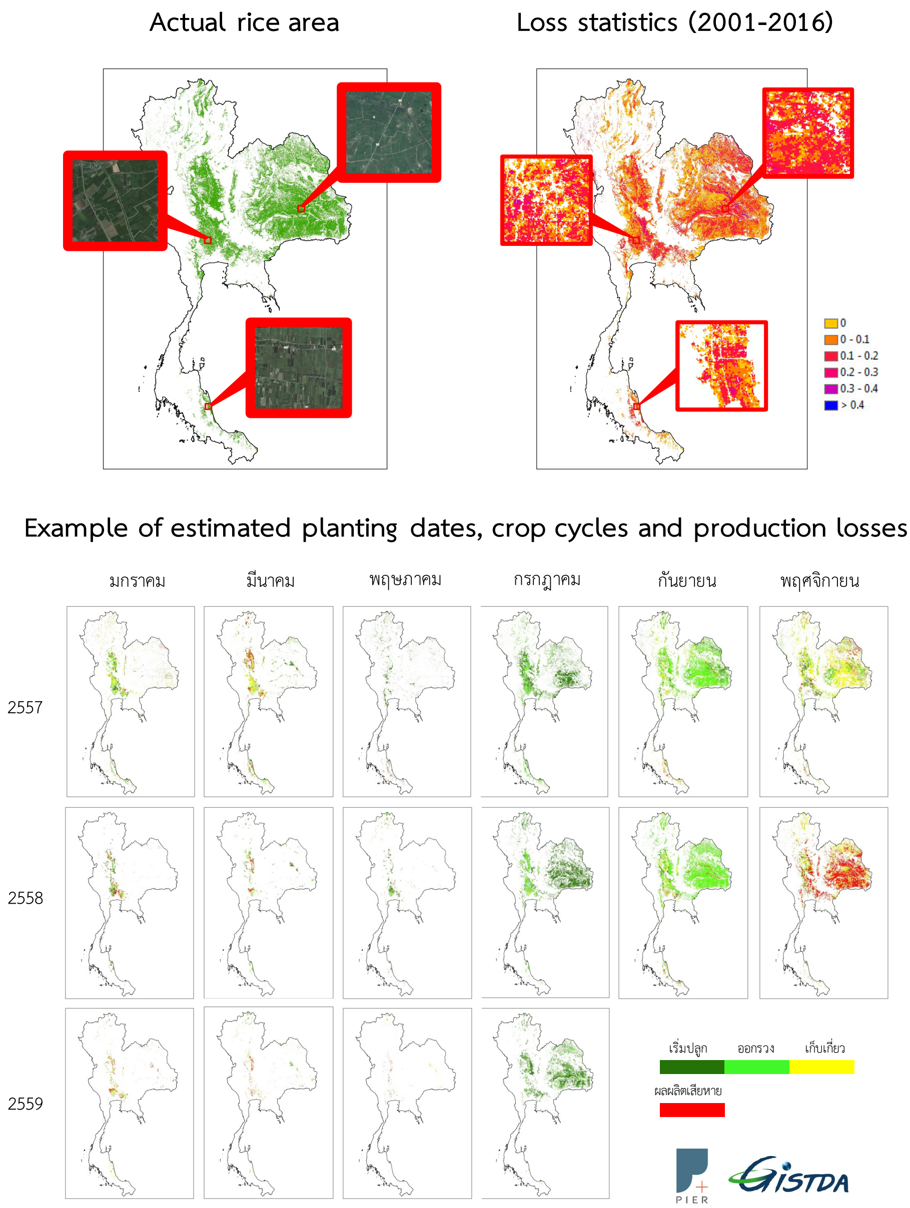


Table 3 Optimal loss parameters

Cluster	Cluster name	Loss threshold (n*) (Percentile)	Total loss threshold (n**) (Percentile)	Loss weight assigned to each growing stage (w_s)				
				Plant 1st month	Grow I 2nd month	Grow II 3rd month	Grow III/Harvest 4th month	Harvest 5th month
1	Central - rainfed	P15	P5	10%	25%	25%	40%	-
2	Central - irrigated - jasmin rice	P15	P5	10%	20%	30%	40%	-
3	Central - irrigated - low variations	P15	P5	10%	20%	30%	40%	-
4	Central - irrigated - high variations	P10	P5	0%	25%	35%	40%	-
5	North - rainfed	P15	P5	10%	20%	30%	40%	-
6	North - irrigated - jasmin rice	P15	P5	10%	25%	25%	40%	-
7	North - irrigated - low variations	P20	P5	10%	20%	30%	40%	-
8	North - irrigated - high variations	P10	P5	10%	30%	30%	30%	-
9	NE - rainfed - jasmin rice	P15	P5	10%	20%	20%	20%	30%
10	NE - rainfed - sticky rice and others	P15	P5	10%	20%	20%	20%	30%
11	NE - irrigated jasmin rice	P20	P5	10%	25%	25%	40%	-
12	NE - irrigated others	P10	P5	10%	20%	30%	40%	-
13	South - rainfed	P20	P5	10%	20%	20%	50%	-
14	South - irrigated - low variations	P15	P5	10%	20%	30%	40%	-
15	South - irrigated - high variations	P10	P5	10%	25%	25%	40%	-

Verification with household data

To what extent could these satellite estimated information reflect actual crop production and losses of farm households? Because the NDVI data are matched with farm households at Tambon level (note that it is also possible to merge at village level, though the process could take some good amount of time due to the lack of village shp file for ArcGIS), this implies that the predictive errors could come from at least two sources: (1) the extent to which the Tambon-level estimated loss can predict Tambon loss and (2) the extent that actual loss of each farm household varies from Tambon loss.

We rewrite farm household i's production loss for crop c, year y and in Tambon T as

$$l_{iy}^{c} = l_{Ty}^{c} + \varepsilon_{iTy}^{c} \tag{9}$$

where l_{Ty}^{c} reflect Tambon-level average loss and ε_{iTy}^{c} reflects variation of household i's loss from Tambon-level average. (9) thus disaggregate household loss into the component that co-move with Tambon-level average and the idiosyncratic component that are specific to each households and uncorrelated with others within Tambon.

We can then write Tambon-level loss as

$$l_{Ty}^{c} = \tilde{L}_{Ty}^{c}\left(n_{pt}^{*}, n_{pt}^{**}, w_s\right) + \varepsilon_{Ty}^{c} \tag{10}$$

And so ε_{Ty}^{c} reflects the extent that the Tambon-level losses could not be explained by satellite-based loss $\tilde{L}_{Ty}^{c}\left(n_{pt}^{*}, n_{pt}^{**}, w_s\right)$. And so from (9) and (10), we can deduce that the total 'basis risk' faced by each farm households from using satellite data to predict loss could thus be decomposed as $\varepsilon_{iy}^{c} = \varepsilon_{iTy}^{c} + \varepsilon_{Ty}^{c}$.

Using 10 years of data of 31,295 farm households, we first estimate the total predictive errors ε_{iy}^{c} and their distributions across different production zones. And Table 4 and Figure 12 depict these results.

We also attempt to understand how predictive errors of satellite-estimated losses can be compared with those from existing field verification from the National Disaster Relief Program. While the data do not come with list of Tambon being declared disaster zone, we rather extract this information from the actual pdf version of announcement of disaster affected areas themselves. Given that the task is very time-consuming, we only manage to do it for 3 consecutive years of 2011-2013. This thus provides additional data on which Tambon in those three years had been declared disaster zones. We can then use this information to analyze the probability that the Tambon was declared disaster zone given that the farm households observed with losses in the household survey. And so in the last column of Table 4, correctly predict loss = 1 if the Tambon that lost household resigned was also declared disaster zone by the National Disaster Relief Program.

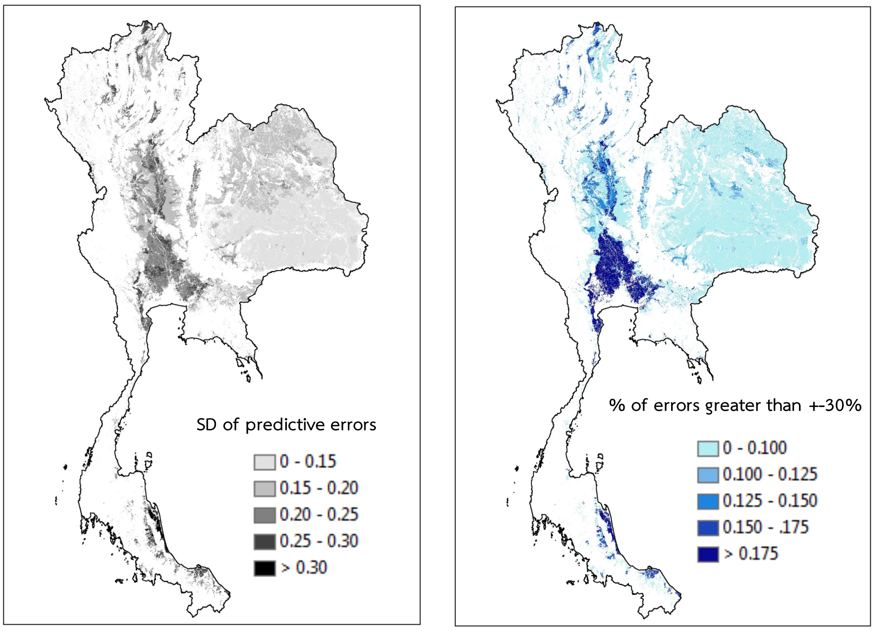
Table 4: Distribution of total predictive errors

Cluster	Cluster name	Satellite-based loss prediction						Current loss verification
		SD of errors	Correctly predict loss (0,1)	Distribution of predictive errors				Correctly predict loss (0,1)
				Within +- 10%	Within +- 20%	Within +- 30%	Greater than +- 30%	
1	Central - rainfed	0.18	79%	29%	45%	18%	8%	77%
2	Central - irrigated - jasmin rice	0.24	66%	26%	28%	28%	18%	75%
3	Central - irrigated - low variations	0.24	69%	24%	29%	26%	21%	81%
4	Central - irrigated - high variations	0.26	57%	19%	35%	23%	23%	79%
5	North - rainfed	0.18	77%	36%	39%	17%	8%	78%
6	North - irrigated - jasmin rice	0.26	68%	25%	34%	25%	16%	77%
7	North - irrigated - low variations	0.23	73%	23%	34%	29%	14%	83%
8	North - irrigated - high variations	0.26	65%	25%	27%	26%	22%	85%
9	NE - rainfed - jasmin rice	0.12	84%	43%	37%	13%	7%	81%
10	NE - rainfed - sticky rice and others	0.15	80%	41%	42%	10%	7%	80%
11	NE - irrigated jasmin rice	0.19	73%	34%	39%	16%	11%	86%
12	NE - irrigated others	0.21	71%	31%	37%	18%	14%	84%
13	South - rainfed	0.25	72%	35%	42%	13%	10%	76%
14	South - irrigated - low variations	0.28	68%	27%	33%	23%	17%	84%
15	South - irrigated - high variations	0.32	59%	25%	36%	18%	21%	79%

We can see that performance of satellite-based estimation of loss vary across production zone with good performance in the rainfed zones especially in the Northeast, Central and North with relatively low standard deviation of errors and 72-84% probability that satellite-based loss can accurately predict the occurrence of loss by household. Performance worsen especially in for irrigated zones in the South, Central and North with relatively larger standard deviation of errors.

Looking at the distribution of predictive errors, we find that especially in the rainfed zones, more than 70% of errors are mainly within +-20% and less than 10% of the time when we see errors beyond +-30%. This patterns also applies to the irrigated zones though we still see quite large percentage of errors in the largest section. Figure 12 plots relative performance of satellite-based loss estimation by production zones.

Figure 12: Distribution of total predictive errors by production zones



Comparing with the loss announced by the National Disaster Relief Program, we find the performance of satellite-based loss estimation to be relatively similar to that of the Relief Program, though the program appears to perform better especially for irrigated areas where there appears to be limitation in satellite estimation.

How might predictive errors differ across different magnitudes of actual losses? Table 5 reports distributions of total predictive errors conditional on actual losses being in one of the three magnitudes. We find that for all clusters, satellite-based loss estimation perform very well in predicting extreme losses of more than 1-in-5 year losses with predictive errors reduce for those loss levels. This could further imply that satellite data could perform well in predicting extreme loss events, while there still needs to add other higher resolution satellite data and to have field verification to improve predictive power in those irrigated areas where production patterns are largely heterogeneous.

Table 5: Distribution of predictive errors conditional on the magnitude of losses

Cluster	Distribution of predictive errors											
	For actual loss < 1-in-3 yr loss				For 1-in-3 yr loss < actual loss < 1-in-5 yr loss				For actual loss > 1-in-5 yr loss			
	Within +- 10%	Within +- 20%	Within +- 30%	Greater than +- 30%	Within +- 10%	Within +- 20%	Within +- 30%	Greater than +- 30%	Within +- 10%	Within +- 20%	Within +- 30%	Greater than +- 30%
1	19%	49%	21%	11%	34%	49%	14%	3%	61%	22%	12%	5%
2	17%	29%	33%	21%	37%	24%	24%	15%	43%	31%	15%	11%
3	21%	22%	32%	25%	25%	40%	19%	16%	35%	38%	14%	13%
4	14%	38%	25%	23%	23%	32%	21%	27%	33%	28%	19%	16%
5	24%	45%	21%	10%	53%	30%	12%	6%	56%	30%	10%	4%
6	25%	23%	32%	20%	12%	60%	17%	11%	47%	34%	11%	8%
7	18%	28%	36%	18%	24%	47%	20%	9%	41%	36%	16%	7%
8	21%	23%	28%	28%	29%	30%	24%	17%	34%	38%	21%	7%
9	28%	49%	14%	9%	65%	19%	12%	5%	67%	19%	11%	3%
10	27%	52%	12%	9%	60%	29%	6%	4%	65%	23%	8%	4%
11	24%	44%	18%	14%	44%	36%	14%	7%	58%	24%	12%	6%
12	22%	38%	22%	18%	38%	42%	11%	9%	55%	25%	14%	6%
13	29%	46%	15%	10%	36%	40%	11%	13%	57%	29%	9%	5%
14	23%	31%	27%	19%	28%	38%	19%	15%	41%	33%	14%	12%
15	22%	32%	22%	24%	24%	50%	13%	17%	38%	28%	11%	15%

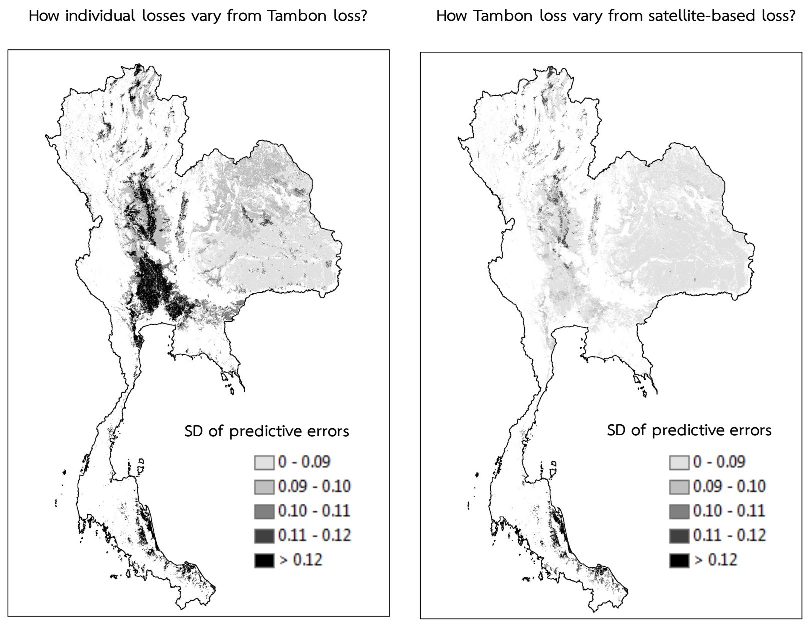
How do we explain these patterns of total predictive errors? We further decompose errors into the errors that reflect how individual losses deviate from Tambon loss ε^{c}_{iTy} (or the idiosyncratic component of loss that varies across households) and those that reflect how Tambon loss deviate from the satellite-estimated loss ε^{c}_{Ty} (or the part of covariate Tambon-level loss that could not be explained by satellite data).

We find in Table 6 that in all production zones, most of the predictive errors are from the idiosyncratic component of losses of each household from the Tambon losses. This could imply that one can improve the total predictive power of satellite data by trying to match them to households at more micro-level, e.g., village level. This could also imply that satellite data appear to perform well in predicting area losses.

Table 6: Decomposition of predictive errors by sources

Cluster	Cluster name	Individual - Tambon avg		Tambon avg - Satellite preduction	
		SD of errors	% of error Greater than +- 30%	SD of errors	% of error Greater than +- 30%
1	Central - rainfed	0.11	5%	0.06	4%
2	Central - irrigated - jasmin rice	0.13	12%	0.09	13%
3	Central - irrigated - low variations	0.14	16%	0.08	14%
4	Central - irrigated - high variations	0.14	15%	0.10	15%
5	North - rainfed	0.09	5%	0.07	7%
6	North - irrigated - jasmin rice	0.13	11%	0.12	13%
7	North - irrigated - low variations	0.15	13%	0.09	11%
8	North - irrigated - high variations	0.16	17%	0.12	13%
9	NE - rainfed - jasmin rice	0.08	4%	0.03	4%
10	NE - rainfed - sticky rice and others	0.09	3%	0.04	4%
11	NE - irrigated jasmin rice	0.12	9%	0.08	10%
12	NE - irrigated others	0.11	11%	0.09	10%
13	South - rainfed	0.13	7%	0.12	9%
14	South - irrigated - low variations	0.15	12%	0.14	13%
15	South - irrigated - high variations	0.17	15%	0.14	14%

Figure 13: Comparison of patterns of predictive errors by sources



Overall, the above exercise thus implies that the moderate resolution MODIS-NDVI when combined with high-resolution RADARSAT flood data has great potential in predicting rice production losses but not everywhere. In particular, satellite data perform well in rainfed production zones and for extreme losses. These implies that there are needs to complement these satellite data with (i) other higher resolution data that could better capture spatial variations of production patterns in the irrigated areas varies across production zones and (ii) rigorous field verification to ensure of accuracy. And since we could only match individual farming household to satellite data at Tambon level, we found the predictive performance is dominated by the variation of individual losses within Tambon. This thus implies that one way to improve performance of satellite data is to match them to household at more micro level, e.g., village.[1]

While this will be to do list for further research, we turn now to illustrate what these satellite-based loss estimations can tell us and how can we make good use of them.

[1] We also try to analyze how MODIS-NDVI can predict planting date and so growing areas at Tambon level. We do so by comparing predicted growing areas each season based on MODIS-NDVI with the data obtained from registered farm households each season obtained from the DOAE. The result appears promising with predictive errors concentrate within +- 20%.

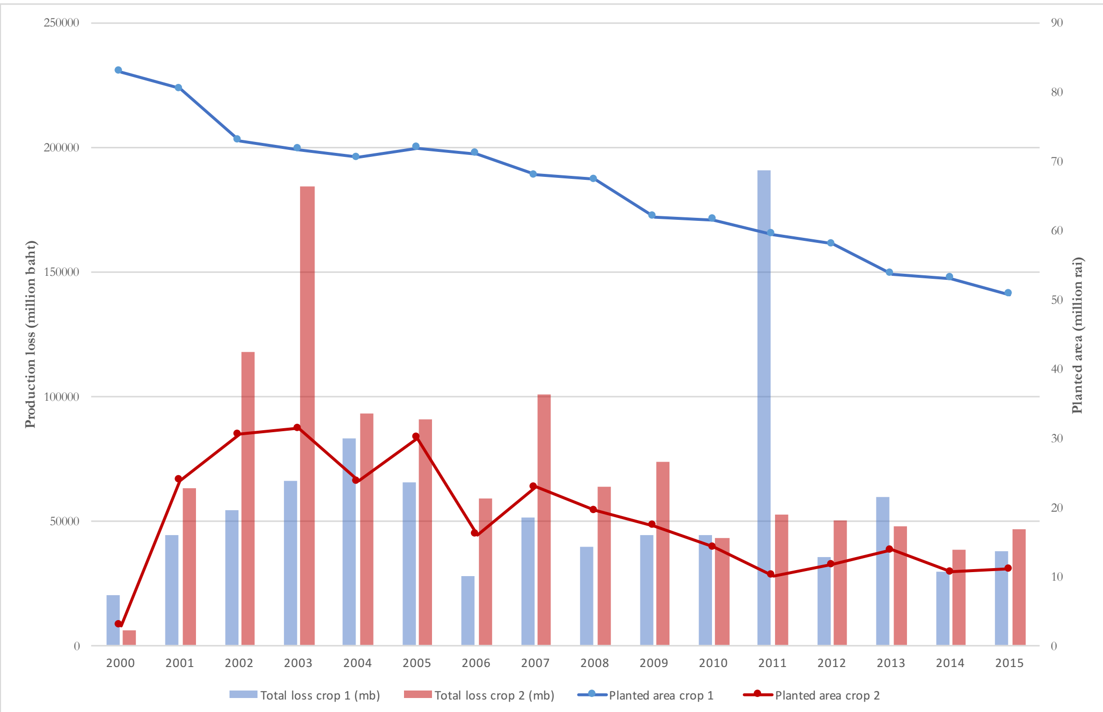
4. How might the satellite-based risk information be used to enhance sustainable agricultural risk management?

What can we learn from this satellite-based risk information?

The constructed satellite-based loss information over time across pixels and rice production areas thus can make very unique and valuable information especially when the data can give information at granular level, over more than 16-year time and over rice growing areas throughout the country. This complements well with other data collected by governments/researchers, which might not be this detailed and with a lot less coverage. This data thus can be used to reflect many different dimension of agricultural risk.

Understanding risk of rice production of the whole country: The very key feature of this satellite-based risk information is the ability to understand dynamics of rice production, losses and risk for the whole country. Figure 14 plots rice production and total losses separated by crop season over the past 16 years. This historical profile gives us big picture of the whole production system of the country, all of which, however are derived from pixel-level information.

Figure 14: Nationwide rice production losses and its distribution



Understanding risk of a farm or an area: More interesting perspectives can be drawn at more micro level. First, with given GIS locator of each farm, satellite-based risk information (or statistics of historical losses) can be extracted for each farm or area, e.g., Tambon or village. This allows us not only to understand cropping patterns, loss over time and statistics of loss – 'risk' – for each farm or area, but also how losses of farms and areas co-move over time to try to understand the extent to which risk can be pooled across geographical areas.

Figure 15: Risk information extracted for each individual farm

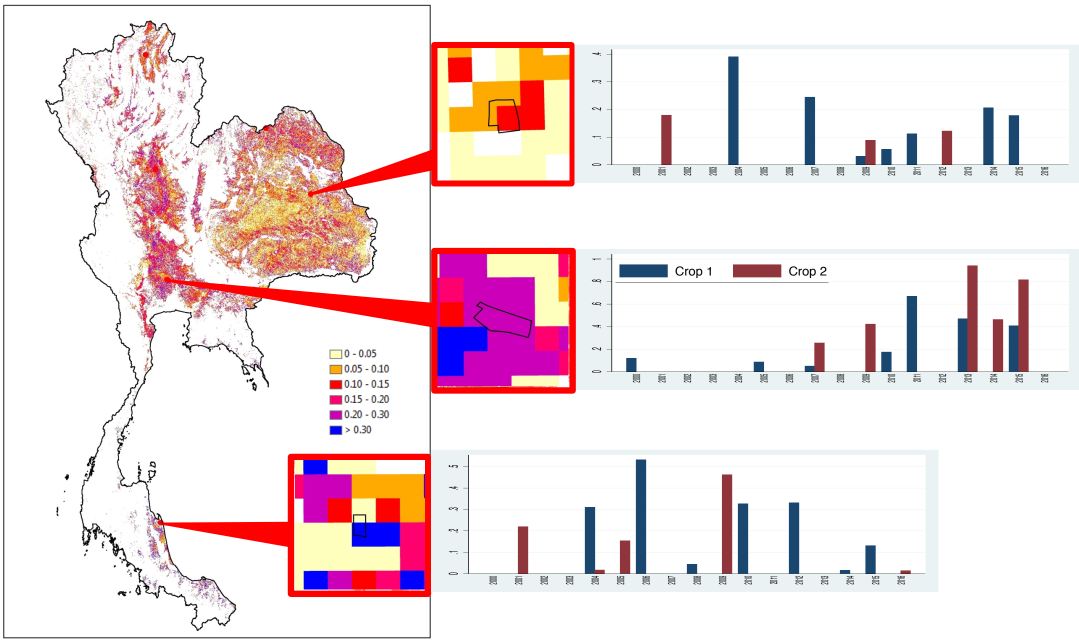
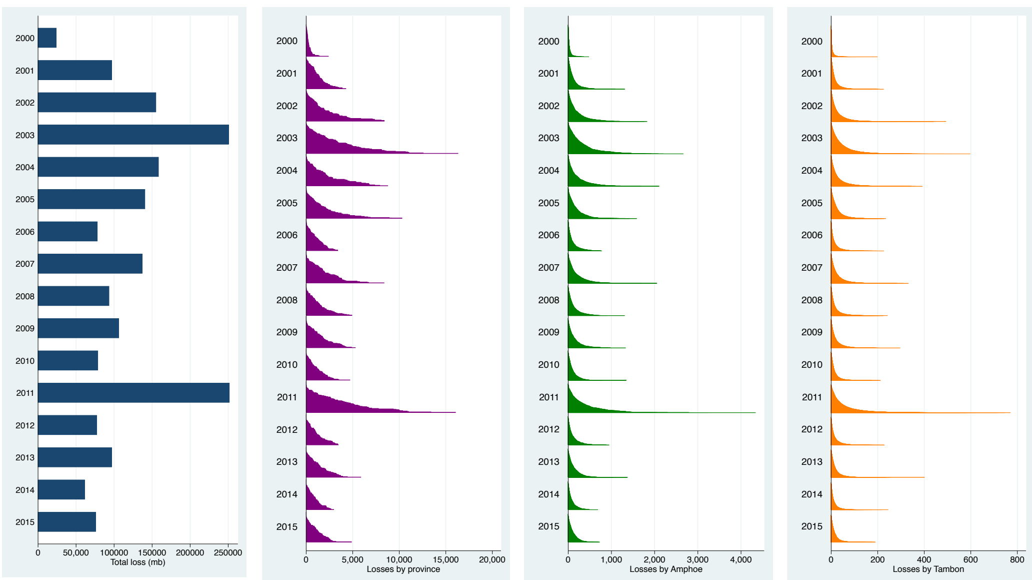


Figure 16: Different level of production losses (million baht)

(a) The whole country



(b) Suphan Buri

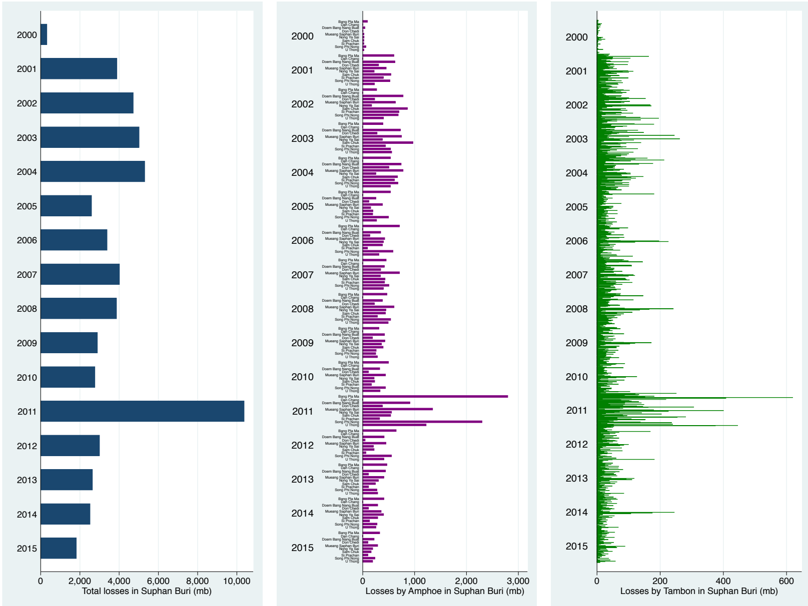


Figure 16 compares losses when measured at aggregated country, province, amphoe and tambon. Different levels give totally different pictures and there are large variations within province and even Amphoe. This implies that using provincial aggregate measures, while could provide big picture of risk, they will not work well in representing loss at more micro level. Losses at Tambon also appears correlated especially in the key flood 2011 and drought 2003 years. This really shows that the rice production losses are largely covariate. There are also some localized events, where losses do not correlate. Figure 17 shows that losses are less correlated across crop seasons but correlated across production zones.

Figure 17: Production losses by crop seasons

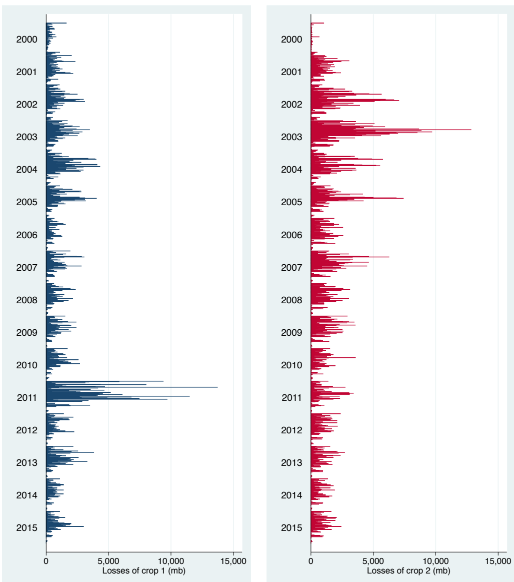
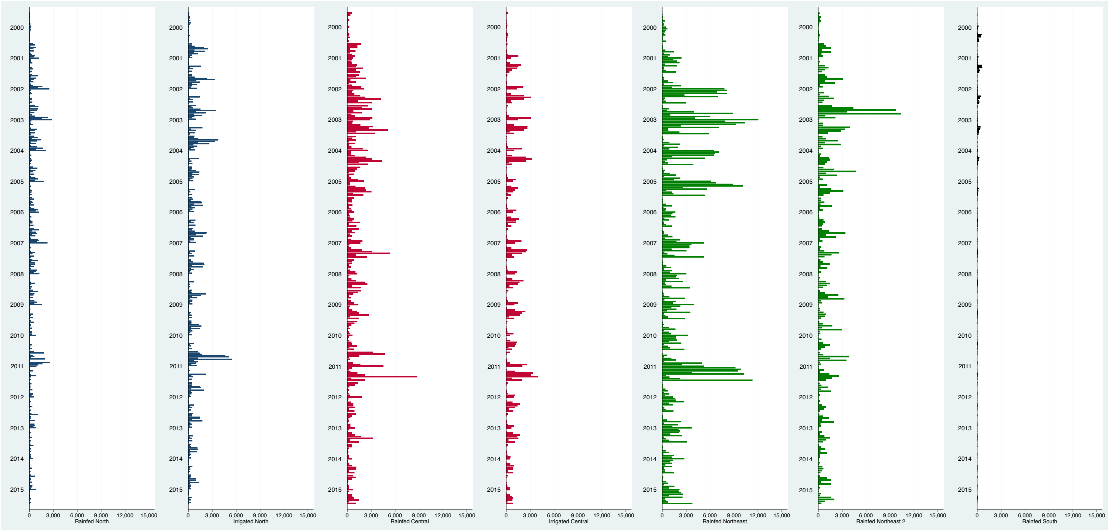
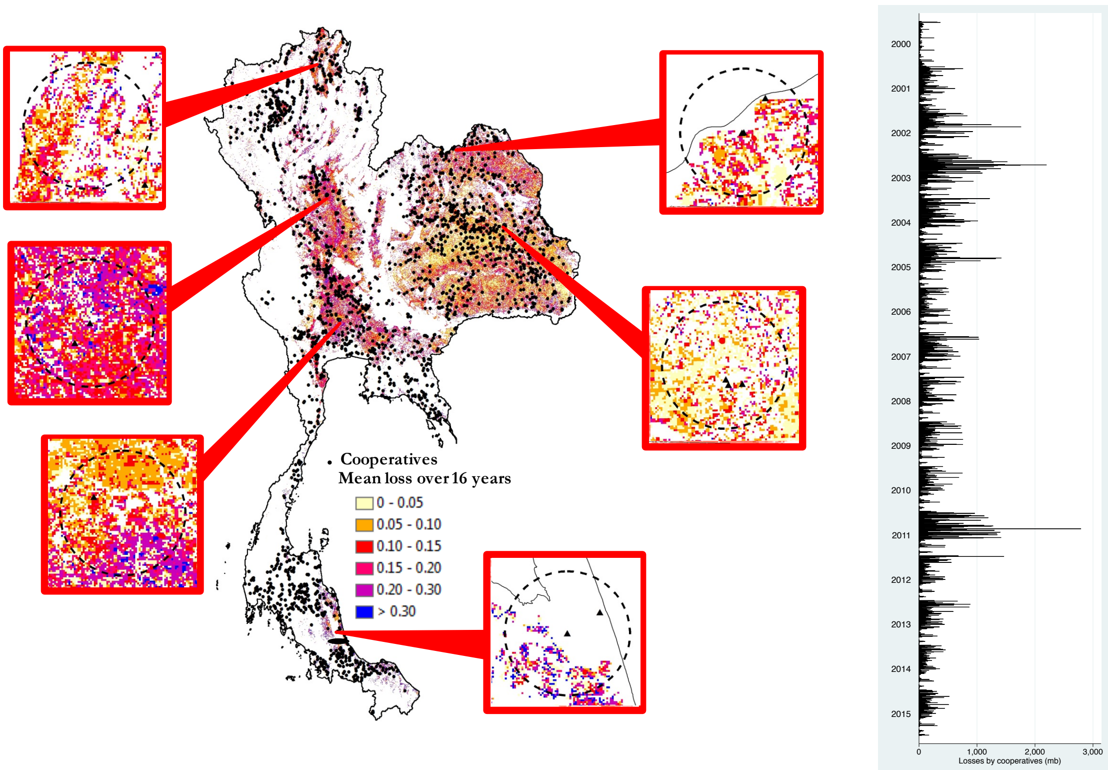


Figure 18: Production losses by selected production zone



Understanding risk of farmers' groups or cooperatives: With location of farmers' cooperatives throughout the country, we can also extract historical losses (i.e., risk profile for each cooperative) as well as profiles of individual farmers belonging to the cooperatives to see how might the patterns of losses vary across households in same period, to what extent might losses co-move or can be shared within cooperatives and to what extent might losses co-move or can be pooled across cooperatives. Figure 18 show that losses appear covariate

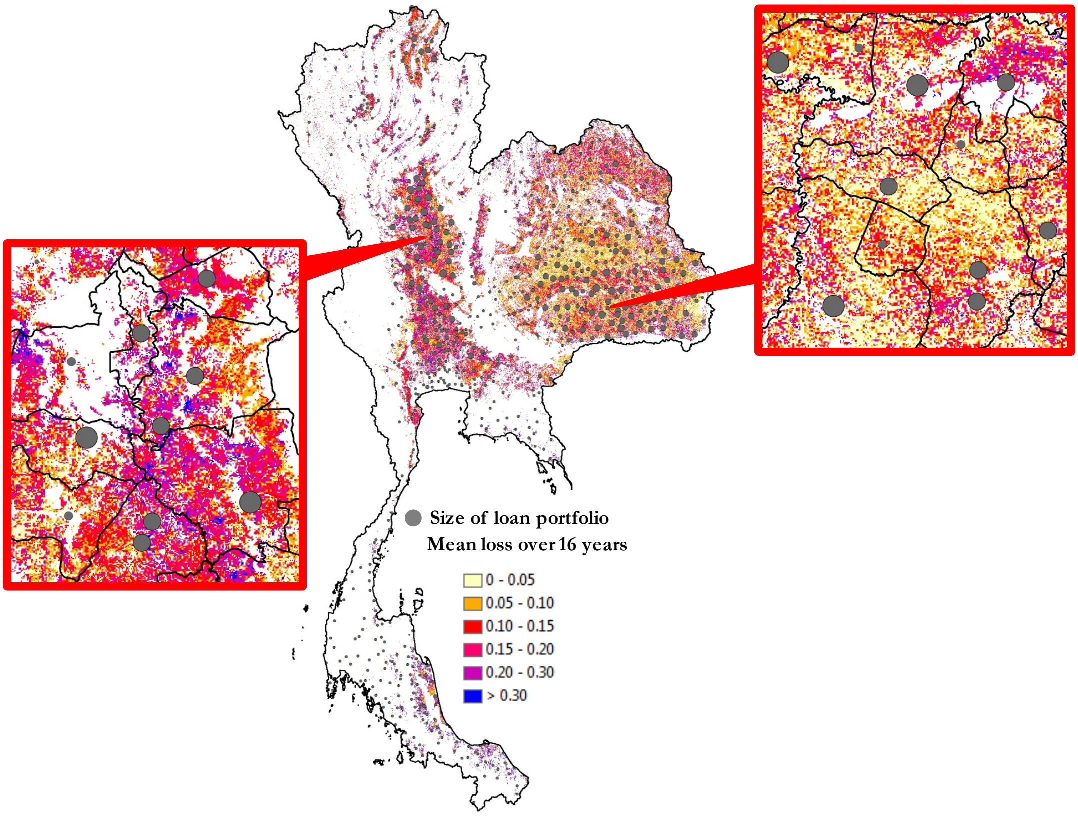
Figure 18: Risk information extracted for cooperative or groups of farmers



across cooperatives. Looking at losses of individual members within each cooperative, we found considerable amount of individual losses that are uncorrelated, implying that there could still be potential for risk sharing across members of cooperatives.

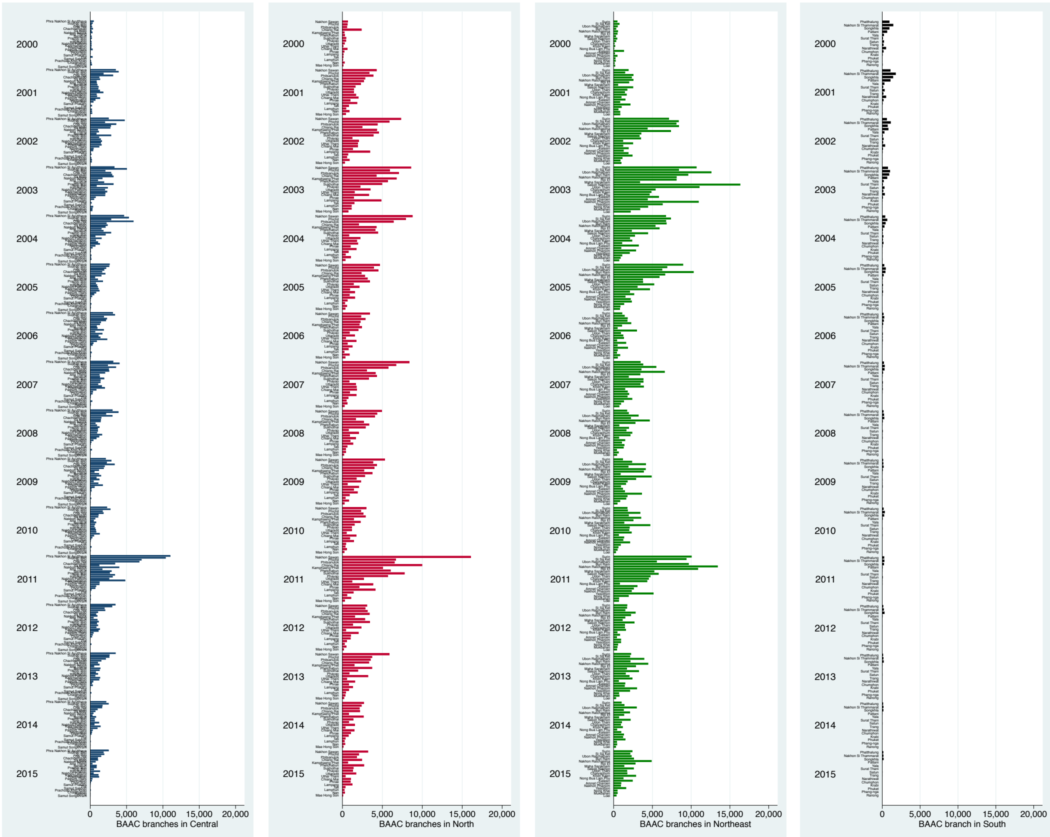
Understanding risk of BAAC's loan portfolio: Figure 19 plots location of all the branches of BAAC along with the size of their loan portfolio to rice farmers in bubble. Risk information can also be extracted for each branch to identify potential credit risk and threat to the bank's overall portfolio. Figure 19 plots rice production losses of farmers belonging to each branch by region. Again, we see high degree of covariate in key disaster years. This results however relies on some assumption on the extent to which credit risk, i.e., the probability that farmers do not repay loan, correlates with their production losses (as we could not obtain any information on actual credit risk from BAAC).

Figure 19: Risk information extracted for each BAAC branch



Production losses (million baht) by provincial branches by region

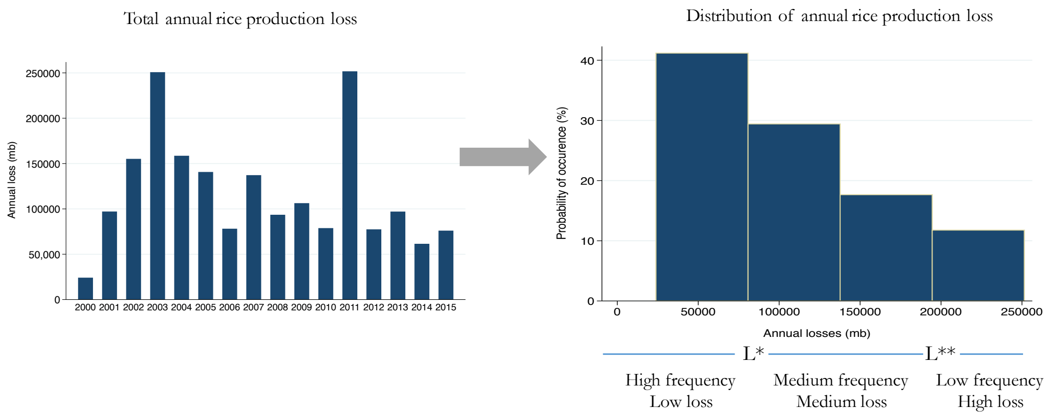
(ranked by total loan outstanding)



Analytical framework

Nature of risk: We first draw empirical distribution of total annual losses for the whole country from 2001-2015 in Figure 20. From Figure 20, we can at least distinguish losses into three layers. First, the losses less than L* reflects high-frequency but low-impact losses, e.g., from localized events that are largely incurred by some farmers and not correlated across farmers, e.g., pest, local flood, etc. Second, the loss from L* to L** reflects in-between losses with moderate frequency and losses. And third, the losses beyond L** reflects low-frequency but highly catastrophic results, e.g., from large disaster events like the flood 2011, etc. Managing these different layers of losses thus would need different costs and for different sectors implying that appropriate risk management tool for each layer could be different.

Figure 20: Distribution of losses and risk layering



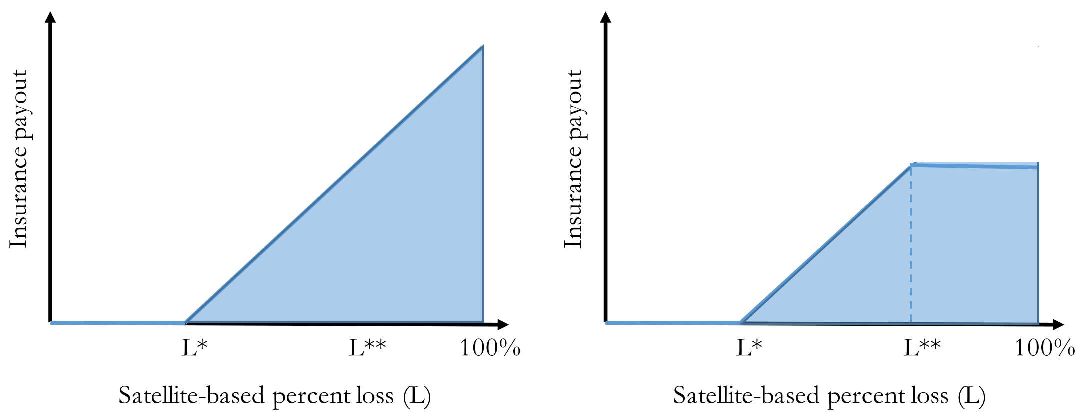
Cost of market risk transfer: Transferring risk to the market requires insurance type financial product that could compensate in the event of losses based on some pre-determined payout function. We consider here the potential to use market risk transfer through insurance contracts, which makes payout based on satellite-based loss data ($\tilde{L}$).

In general, payout function can be illustrated in Figure 21, where on the left, payouts will be made if satellite-based losses is beyond L* (strike level). The right panel applies a threshold L** (exit level) so that payout is equal to maximum payouts if satellite-based losses are beyond L**. In particular, we can write a prototype insurance payout as

$$\pi(\tilde{L}, L^{*}, L^{**}) = \min\left(\max\left(\tilde{L} - L^{*}, 0\right), L^{**} - L^{*}\right) \tag{11}$$

And so the strike L* thus will reflect coverage of the contract, the higher the L* the less coverage this insurance is. The exit L** also reflects the reverse coverage. High L** implies larger coverage of contract. The left panel of Figure 21 has L** at 100% and hence would have more coverage than the one on the write.

Figure 21: The two payout functions considered in this study



Actuarially fair rate for any insurance contract that represents fair pricing just enough that the insurer will cover all the cost of payout is simply the expected value of payouts made based on historical data and hence will be some function of risk information and coverage and so equal to $E\pi(\tilde{L}, L^{*}, L^{**})$. Commercial premium rate is defined as

$$P\pi = catload \times E\pi\left(\tilde{L}, L^{*}, L^{**}\right) + adminload. \tag{12}$$

Where $catload$ is catastrophic load to capture capital cost needed in order to manage some reserve to prepare for times when insurers will need large leverage and so $catload = c(VAR(\pi), quality\ of\ information)$ where $VAR(\pi)$ is value at risk of portfolio payout reflecting probability of extreme events. Catastrophic load will be large with large $VAR(\pi)$ and large when low quality risk information as insurers will then need to prepare for the possibility of events that could be underestimated by low-quality risk information. The influence of bad information to risk transfer cost is amplified when the contract can have extreme payouts as insurers will incur expenses to prepare for these extreme events in the presence of low quality information. The $adminload$ reflects some fixed and variable costs of developing, designing, marketing and selling the contract.

So overall, we can induct that the cost of risk transfer through any insurance contract $P\pi = p(\tilde{L}, L^{*}, L^{**}, VAR(\pi), quality\ of\ information)$. And so the cost will be high with larger coverage (small L^{*}, large L^{**}), larger possibility of extreme payout events and lower quality of risk information. We note that the very key determinant of $VAR(\pi)$ is the potential for risk diversification e.g, through large scale of operation and larger geographical and product coverage.

Farmer's optimal insurance decision: An optimal insurance design defines a combination of coverage (L^{*}, L^{**}) that maximizes the farmer's welfare given the pricing structure. For simplicity, we consider a risk averse household with preference over consumption represented by a class of mean variance utility function with $\theta > 0$ representing an Arrow-Pratt coefficient of absolute risk aversion.[2] With stochastic net income from rice production y_{it} under assumed deterministic price, income from other source yo_{it}, insured household's income available for consumption c_{it} can thus be written as

$$c_{it} = E(y_{it})(1-l) + I_i \times \left(\pi\left(\tilde{L}, L^{*}, L^{**}\right) - P\pi\left(\tilde{L}, L^{*}, L^{**}\right)\right) + yo_{it} + a(\bar{L}) + \gamma_{it} \tag{13}$$

[2] In specific, we consider CARA utility function $U(c_{ilt}) = -e^{-\theta c_{ilt}}$ so that $EU(c_{ilt}) = E\left(-e^{-\theta c_{ilt}}\right) = E(c_{ilt}) - \frac{\theta}{2}\sigma^2(c_{ilt})$ under normally distributed consumption (Ljungqvist and Sargent 2004).

where l reflects farm household's actual rice production losses, I reflects household's insurance decision in term of total amount of rai and sum insured per rai (i.e., value per rai to be covered by insurance, e.g., total income or input cost) they choose to insure, $a(\bar{L})$ reflects other public assistance based on some loss verification $(\bar{L})$ e.g., the current disaster relief and γ_{it} reflects other (potentially positive) stochastic component of rice production income that could not be captured through the expected rice income and losses. Subjected to (11), (12) and (13), an optimal insurance contract for $\pi(\tilde{L}, L^*, L^{**})$ designed based on satellite-based loss estimation $\tilde{L}$ involve searching for a combination of insurance decision I_i^* and contract coverage L^*, L^{**} such that

$$\max_{I_i^*, L^*, L^{**}} E(c_{it}) - \frac{\theta}{2}\sigma^2(c_{it}) \tag{14}$$

Using household data from rice production area nationwide, this allows us to search for optimal coverage level L^*, L^{**} for each rice production zone and individual insurance decision $I_i^* = I(\theta, cov(l, \tilde{L}), L^*, L^{**}, P\pi(\tilde{L}, L^*, L^{**}), cov(l, a(\bar{L})), cov(l, yo_{it}))$. This implies that farmer's optimal insurance decision will be increasing in risk aversion, how satellite-based loss co-move with actual farm loss, coverage level and if farmer has other sources of income that tend to co-move with rice production and so providing less diversification to their overall income. The insurance decision will be decreasing in commercial cost of risk transfer and if the amount of public disaster assistance co-move well with actual loss.

We simulate 1,000 representative farm households from each rice production zone, their rice production income, losses and other income from 2001-2016 based on the estimated relationship with 2001-2016 satellite-based loss estimation in (9) and (10) and parameters observed in the current survey on Table 2. Detail of parameters used and simulation process are reported in Appendix. This allows us to search for the optimal contract coverage by production zone as well as the optimal insurance decision for each and every 15,000 farmers. This allows us to understand insurance demand and its geographical distribution $f(I)$ conditional on the optimal insurance contract coverage.

Government's contribution to risk management: Government's contribution to risk management could come in the form either of disaster relief $a(\bar{L})$ based on some public loss verification $\bar{L}$ and supports to development of insurance market either in the form of premium subsidy $s\left(f(I), P\pi(\tilde{L}, L^*, L^{**})\right)$ and/or acting as insurer of last resort or providing reinsurance s' (either by insuring the extreme losses of portfolio of insurer when portfolio loss or providing direct payouts to insured farmers instead of market beyond the point where it might be too expensive for market to offer (e.g., above L^{**}) and so $s'\big(f(I), \tilde{L}, L^{**}\big)$. And so the total cost to government in risk management can be written for some cost structure $g(\cdot)$ which could reflect both capital cost and opportunity cost of budget as

$$G_t(a(\bar{L}), s, s') = g\left(a(\bar{L}) + s\left(f(I), P\pi\big(\tilde{L}, L^*, L^{**}\big)\right) + s'\big(f(I), \tilde{L}, L^{**}\big)\right) \tag{15}$$

The cost of disaster relief $a(\bar{L})$ will thus be increasing with the farm losses. The cost of subsidy s will be increasing in the amount of farmers insured and commercial pricing of insurance, which then depend on the quality of data and probability of extreme risk, scale and risk diversification potential of the program. And the cost of reinsurance support s' will increase with probability of extreme risk and the amount of farmers insured.

The framework above implies different impacts on farmers of government supports to risk management. Provision of disaster relief, depending on $cov(l, a(\bar{L}))$ will reduce risk of farmers and with high covariance potentially reduce demand for insurance. So well perform disaster relief will result in government acting as insurer for farmers destroying the sustainability of insurance market.

The premium subsidy $s\left(f(I), P\pi\big(\tilde{L}, L^*, L^{**}\big)\right)$ will bring down commercial premium faced by farmers and so could potentially increase demand. But in the setting with probability of extreme payout events in place as well as low quality data, resulting in extremely high premium pre-subsidy, this could imply that government's subsidy might need to be very large in order to bring down affordable premium to farmers or else the subsidy might not work well in stimulating demand. The potential impacts of subsidy that could not capture in our model is the extent to which farmers will depend on subsidy and so the market might not be able to work on its own without government's provision of subsidy.

Government's contribution by acting as insurer providing reinsurance $s'\big(f(I), \tilde{L}, L^{**}\big)$ could bring down commercial premium by reducing coverage from insurers on extreme risk, reduce the impacts of low quality data on commercial premium and so increase demand. Overall, with increasing scale and could further bring down premium. This support appears quite similar to those of disaster relief but the key different is the more targeted losses that government will help and the clear link to market-based insurance (as farmers will have to buy insurance first in order to be quality for this assistance) allowing this support to further stimulate (rather impeding) demand for insurance. We estimate the cost to government applying some cost of capital r.

With government supports, farm households' income available for consumption thus can change to

$$c_{it}^{g} = E(y_{it})(1-l) + I_i \times \left(\pi\left(\tilde{L}, L^{*}, L^{**}, s'\right) - P\pi\left(\tilde{L}, L^{*}, L^{**}, s, s'\right)\right) + a(\bar{L}) + yo_{it} + \gamma_{it} \tag{16}$$

Optimal Public-Private Partnership (PPP) arrangement in risk management: We then define the optimal combination of contribution to risk management from farmers I^{*}, the optimal market-based contract coverage L^{*}, L^{**} and government G^{*} that maximizes welfare of farmers following (14) such that

$$\max_{I_i^{*}, L^{*}, L^{**}, G^{*}} E(c_{it}) - \frac{\theta}{2}\sigma^{2}(c_{it}) \tag{17}$$

Once again, we estimate this using 15,000 simulated farm household data from 2001-2016, where we consider three different types of insurance.

Optimal risk strategies with satellite-based risk information

A typical crop insurance for farmers: The simple payout function can be defined as

$$\pi^{farmer}\left(\tilde{L}, L^{*}, L^{**}\right) = \min\left(\max\left(\tilde{L} - L^{*}, 0\right), L^{**} - L^{*}\right) \times X \times I \tag{18}$$

where insurance payout is based on observed satellite-based losses $\tilde{L}$ measured at Tambon level where farmer locates relative to the strike and exit, X reflect sum insured of the contract. The insurance can be written on crop income with predetermined sum insured, e.g., $X = \bar{p}\bar{y}$ where $\bar{p}$ reflects some pre-determined average price per kg, $\bar{y}$ reflects pre-determined average yield per rai in kg and so I here reflects the number of rai that farmers would like to insure. The commercial premium thus can be quoted as percentage of total sum insured, i.e., as percent of $X \times I$.

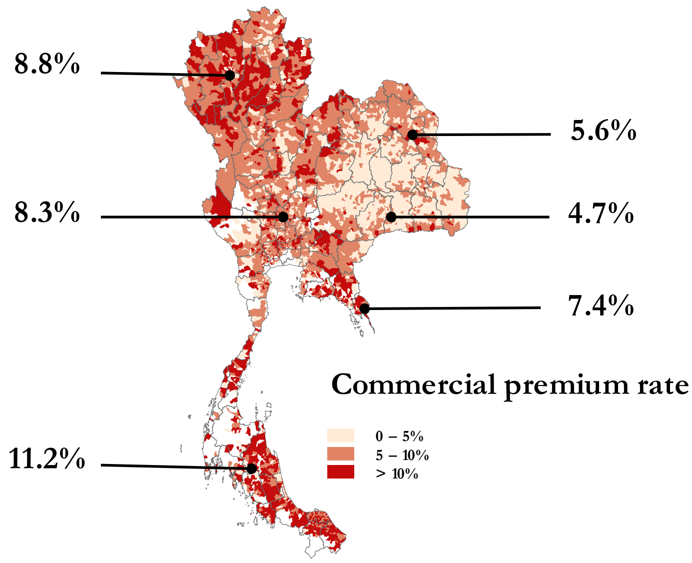
Table 7 presents optimal PPP arrangements for each production zone and how the optimal arrangement might change with changing insurance scale and coverage. We find that the optimal insurance coverage levels vary across production zones and range in between 1-in-4 yr and 1-in-6 yr losses. We find that the optimal government support is to pull resource to providing assistance (i.e., as insurer of last resort) for extreme losses that result in more than 70% of production losses of the country, which occur 1-in-10 year in the history in 2011 flood and 2015 drought. Farmer's demand for insurance varies across production zones with high demand in the zones where satellite-based loss explains actual farmer's loss well, e.g., in the rainfed areas. High demand also appears in the very risky irrigated areas. And lastly, using various risk aversion parameters to simulate results, we find that insurance demand increases with risk aversion of farmers.

We also estimate these results in the case when only crop 1 is insured, the result (not shown here) also reflects the possibility of risk diversification when contracts extend coverage throughout the country and to cover both crop 1 and crop 2 seasons. And so overall, with optimal government's support that can complement (not substitute) market, we find increasing market participation by farmers, which could then increase scale and risk diversification and hence sustainability of the system.

Table 7: Optimal PPP arrangements with market-based crop insurance to farmers

Cluster	Cluster name	Market		Government		% Farmer demand	
		L*	Commercial premium rate (%)	Subsidy (g)	Insurer of last resort (g'=L**)	Low risk aversion (theta <3)	High risk aversion (theta >3)
1	Central - rainfed	30.0%	8.4%	0.0%	60.0%	59.0%	78.0%
2	Central - irrigated - jasmin rice	30.0%	6.4%	0.0%	60.0%	61.0%	84.9%
3	Central - irrigated - low variations	30.0%	7.3%	0.0%	60.0%	67.0%	97.5%
4	Central - irrigated - high variations	25.0%	6.3%	0.0%	60.0%	56.0%	95.8%
5	North - rainfed	30.0%	8.1%	0.0%	60.0%	54.0%	71.0%
6	North - irrigated - jasmin rice	35.0%	9.0%	0.0%	70.0%	43.0%	67.0%
7	North - irrigated - low variations	30.0%	7.7%	0.0%	60.0%	46.0%	69.5%
8	North - irrigated - high variations	25.0%	7.0%	0.0%	60.0%	47.0%	65.0%
9	NE - rainfed - jasmin rice	30.0%	5.1%	0.0%	60.0%	61.0%	77.0%
10	NE - rainfed - sticky rice and others	30.0%	6.9%	0.0%	60.0%	57.0%	75.0%
11	NE - irrigated jasmin rice	30.0%	7.4%	0.0%	60.0%	43.0%	67.0%
12	NE - irrigated others	30.0%	4.7%	0.0%	60.0%	32.0%	53.0%
13	South - rainfed	40.0%	8.5%	0.0%	70.0%	28.0%	45.0%
14	South - irrigated - low variations	25.0%	9.1%	0.0%	60.0%	25.0%	34.0%
15	South - irrigated - high variations	30.0%	8.2%	0.0%	60.0%	21.0%	39.0%
All	Nationwide	-	6.5%	-	-	45.0%	63.0%

Figure 22: risk-based pricing of optimal crop insurance to farmers (by Tambon)



A group insurance for cooperatives or farmers' groups: The key different and potential advantage of group-based insurance comparing to the insurance to farmer are (1) the possibility of group risk sharing could allow farmers to share basis risk (when $\tilde{L}$ does not perfectly reflect actual loss l of farmers) and so reduce the adverse impact of basis risk if basis risk across farmers within group are uncorrelated (hence increase value of insurance to farmer and so increase demand) and (2) the possibility of risk pooling across Tambon if cooperatives are large enough and losses are less correlated across Tambons (hence reduce insurance price and so increase demand).

And so the payout for group insurance can simply be written as

$$\pi^{group}(\tilde{L}, L^*, L^{**}) = \min(\max(\tilde{L} - L^*, 0), L^{**} - L^*) \times X \times I \tag{19}$$

where X now reflects sum insured of the whole group, e.g., the group-insurance pool received from individual contribution of member, etc.

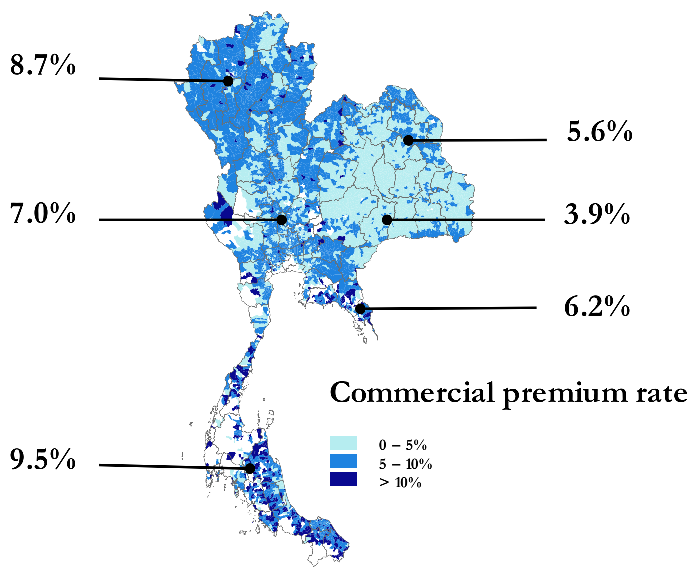
The working assumption for this group insurance is the existing of group-based risk sharing within cooperative so that following (9) and (10), farmer's loss after sharing within group is

$$l_{it}^{shared} = l_{groupt} = \frac{\sum_{i\,in\,group} l_{it}}{\sum_{i\,in\,group} i} = \tilde{L}_t + \frac{\sum_{i\,in\,group} \varepsilon_{Tt}}{\sum_{i\,in\,group} i} + \frac{\sum_{i\,in\,group} \varepsilon_{it}}{\sum_{i\,in\,group} i} \tag{20}$$

where $\tilde{L}_t$ reflect satellite-based losses at Tambon where farmer locates, ε_{Tt} is variation of tambon loss from satellite-based loss and ε_{it} is variation of individual loss from tambon loss. And so the reduction in basis risk will increase if the two sources of basis risk can be shared or as $\frac{\sum_{i\ in\ group} \varepsilon_{Tt}}{\sum_{i\ in\ group} i} \to 0$ and $\frac{\sum_{i\ in\ group} \varepsilon_{it}}{\sum_{i\ in\ group} i} \to 0$.

Table 8 presents optimal PPP arrangements if group insurance can be used for each production zone and how the optimal arrangement might change with changing insurance scale and coverage. We can see that with group-risk sharing in place, the optimal group insurance coverages are lower than those of insurance to farmers as the group insurance can be used to complement group-risk sharing within cooperative, which already takes care of most of the localized and idiosyncratic risk. The key improvement from individual contract comes from larger market participation, which is due to reducing in basis risk, which now can be shared within cooperatives, and not from the reducing premium due to risk sharing across Tambons in the cooperative as Tambon losses are highly correlated even across Tambons within the same neighborhood. This group insurance appears as very promising model.

Figure 23: Risk-based pricing of optimal group insurance to cooperatives



A risk contingent credit for farmers: Apart from using insurance to insure farmers or group of farmers' crop income, farmers also take out input loan at the beginning of the cropping season and the production losses could directly affect ability to repay loan and so to accumulate debt of farmers. The BAAC could also experience reducing loan portfolio quality when rice production losses of borrowers are largely stochastic.

Agricultural banks worldwide offer risk contingent credit to farmer, which really is the loan that is linked with insurance. Farmers take out loan paying interest which will be slightly higher due to the inclusion of insurance premium that covers the total loan plus interest. In the normal years, farmer needs to repay the full loan but in the bad years, when farmers experience production losses, they could just need to repay partial of the loan and the rest will be repaid by the insurance company directly to the bank. This contract thus works as the bank obtaining insurance for their overall loan portfolios, of which they then pass on burden of insurance premium to borrowers.

With limit information on loan performance of BAAC, we assume for now that the loan portfolio losses co-move with the rice production losses with correlation δ. And so total losses of loan portfolio of BAAC proxy by satellite data can be written as

$$\tilde{L}_t^{loanport} = \sum_{branch} outstanding_{bt} \times \tilde{L}_{bt} = \sum_{branch} outstanding_{bt} \times \delta \times \tilde{L}_{at} \quad (21)$$

where $\tilde{L}_{bt}$ reflects loss of loan from branch b in year t, which is calculated from the satellite-based loss of the Amphoe for that branch. And so insurance that can protect loan portfolio of BAAC can be written as

$$\pi^{loanport}(\tilde{L}, L^{*}, L^{**}) = \min\left(\max\left(\tilde{L}_t^{loanport} - L^{*}, 0\right), L^{**} - L^{*}\right) \times X \times I \quad (22)$$

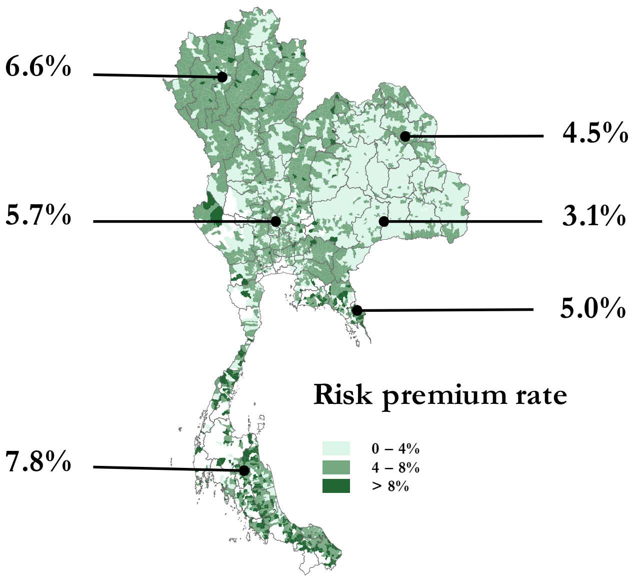
And so for farmer i, taking risk contingent loan means that they will pay slightly higher interest on their loan with return of the following loan repayment function

$$LR_{it} = (1 + r + p\pi) loan - \min\left(\max\left(\tilde{L}_t^{loanport} - L^{*}, 0\right), L^{**} - L^{*}\right) loan \quad (23)$$

And so the key benefit of risk contingent loan is to protect farmers from being indebtness, protect BAAC portfolio and if implemented widely for BAAC's borrowers, this could benefit from the scale of contract and risk diversification from the nationwide rice production loan portfolio.

Table 9 report results the optimal PPP arrangements when risk contingent loan is in place. With results on optimal contract similar to insurance for farmers, we can move to construct interest rate on this risk contingent loans conditional on the assumed degree to which credit risk correlate with production risk. This interest based is closed to risk-based pricing of loan when $\delta \rightarrow 1$.

Figure 24: Risk premium of optimal group insurance to cooperatives



How might these satellite-based risk management strategies work?

So far we already illustrate the potential that satellite-based risk data can be used to provide necessary risk information to allow government, market and farmers to contribute to the optimal PPP arrangement of risk management, which involves farmer's lively contribution to market-based insurance and government acting as insurer of last resort providing assistance to insurer or directly to insured farmers when extreme events occur. How might this system work in the real world?

First farmers have to register their farm and production with the DOAE and through the newly establish GISAGRO platform, they can also draw GIS locator of their farm allowing production cycle, loss and risk to be detected by satellite data linking them to easily by farm location. This individual farm information is then incorporated with satellite data as a center risk information system, which could then provide high quality risk information for insurers, BAAC and government.

The insurers can then use to design various insurance products and sell at risk-based pricing in the market. They can use satellite data (perhaps together with field-based verification) to identify eligibility of farmers (whether they really grow rice in the plots, is growing rice during the insured season), losses, and how they might change growing patterns following getting insurance.

BAAC can also use risk information to monitor their credit risk to loan portfolio, to design various risk contingent credit contracts to protect themselves and farmers, allowing them to offer insured loan to farmers with slightly higher interest in exchange for some protection of loan in bad years. Government can also use information to improve their assistance using satellite-based loss as basis for deciding assistance in the extreme years. This satellite-based risk information comes in near real time thus allowing faster assistance to be implemented.

Farmers thus can have wide range of insurance and credit products available to them at reasonable prices, which could be sold through BAAC and link with the BAAC account. This includes income insurance to protect crop income and insured loan to protect input loan to BAAC. When group insurance is now available, this could perhaps strengthen group-risk sharing mechanisms, where farmers can regularly contribute payment to the group pool of which they can take out as loan (or cash) in case they experience losses. The group can insure the pool so that when everyone in the group experience losses at the same time so the pool run out, insurance can provide payout to leverage it. Farmers can also use risk information to make production decision, stocking and harvesting to control exposure to over-supply prices.

5. Potential economic impacts

Finally, we simulate the potential economic impacts of the satellite-based PPP arrangements in risk management and comparing with existing program where starting from this year government will buy insurance for all the rice farmers who are borrowers of BAAC and use current government loss verification to trigger payout at the commercial cost of 100 baht per rai. The results are based on 15,000 simulated farm household data from all production zones throughout the country. In All the simulated model and parameters are reported in the Appendix.

We note however that this comparison between the two systems has not hold anything equal given the difficulty and large difference across the two. The very key difference are

(1) The current system is no longer commercial based as government is the sold buyer and so government would end up spending large amount of money both for insurance premium for all farmers and for disaster relief. This really is the case as government insure

part of their budget exposure to providing relief to farmers through market and take on the rest of the exposure themselves. This in contrast to our proposed model where government rather acts as insurer of last report providing assistance only for the extreme risk, where it is not cost effective for market to do so.

(2) Insurance coverage will be a lot different with the current program providing payout only when total loss occurs. The current program covers 30% of input cost as part of free insurance and another 30% from relief. And so for farmers, they can be protected up to 60% of input cost only in the bad years. Our insurance is designed to compensate income losses and another program to compensate for all input loans. So farmers will be protected both from the income and input cost that really affects their ability to repay loan and income available for consumption.

Based on 15,000 simulated farmers over the 16 years when we have satellite data, we simulate the potential impacts on (i) farmer's income available for consumption and their probability of defaulting on input loan, (ii) proportion of defaulted input loan of BAAC loan portfolio to rice farmers, (iii) government budget exposure. These simulations aim to compare the impacts of the current program and (1) satellite-based PPP arrangement when crop income insurance to farmers are available, (2) satellite-based PPP arrangement when crop income insurance to cooperatives are available for existing cooperatives and (3) satellite-based PPP arrangement when contingent loan are available and (4) PPP arrangement when both crop income insurance and contingent loan are available.

Impacts on farm households: We find that while current program can reduce probability of farm income available for consumption dropping toward zero, there are still some good probability of low consumption, which could result from low insurance coverage. The satellite-based crop income insurance can eliminate such low consumption probability especially when both crop income insurance and risk contingent credit are implemented. The optimal satellite-based PPP arrangements thus result in lower debt accumulation to farmers relative to the current program. These are average results, we however find impacts also vary across production zones and lower for irrigated zones where satellite-based loss estimation are less correlated with actual losses (results are not shown here).

Impacts on BAAC loan portfolio: With the assumption that farmers will always pay back loan as much as they can with positive income available for consumption after meeting subsistence consumption level at the poverty line (which could already be too optimistic), we find probability of default decreases substantially with the satellite-based program. This is contrasting to the current program with low insurance coverage.

Impacts on government's budget: We find that the current program exposes government with large budget exposure having to both provide disaster relief and pay for the insurance premium. When government 's support is optimized and confined to only providing assistance during extreme losses and so allowing market to work properly and farmers to insure themselves, these result in less budget exposure.

Impacts on insurer's profitability: Would the proposed satellite-based risk management system profitable for the private insurers and so could enhance their incentives to step in to using this improved quality risk information to innovate and offer more variety of financial contracts to farmers? We simulate loss ratio of the insurers' nationwide rice insurance portfolio when risk are pooled for 2, 5 and 10 years. We find that while there could still be possibility of some losses for the first couple years of operation, such probability reduces substantially when the risk are pooled beyond 5 years. Of course, we have not consider the potential risk pooling across portfolios of other types of risk.

Finally, since we have not hold the government budget spent on each program constant when doing the comparison. For comparison, we estimate the expected outcome for 1,000 million baht that government expects to spend. And Table 10 reports this result. We find that various potential insurance and PPP arrangements that are brought about by satellite-based risk information outperform existing program in terms of decreasing crop income variability and probability of defaulting input loan of rice farmers as well as in term of decreasing proportion of defaulted loan in BAAC loan portfolio.

Table 10: The potential impacts on farmers, BAAC and governments

Outcome	Existing program	Crop income insurance	Cooperativ e insurance and PPP	Insured loan and PPP	Crop insurance and insured
Decrease farmer's income variabiility	2.8%	11.2%	13.0%	3.6%	15.8%
Decrease farmer's prob. of loan default	1.8%	6.4%	8.8%	11.4%	13.5%
Decrease BAAC's portfolio loan default	1.4%	6.2%	7.6%	12.4%	13.6%

These simulated results thus make a good evidence for the potential economic values to investing in risk information, one of which can be satellite data.

6. Conclusions and policy implications

This paper already illuminates the promises of satellite data (that are collected transparently, continuously and near real time at granular level but covers nationwide and more than 16 years) in creating high quality risk information necessary to unlock development of sustainable risk management in agriculture. We show however that despite great performance in most of the areas, there are room for improvement especially in highly heterogeneous production zones in irrigated cropping areas, implying the need to invest in and explore the use of better resolution satellite data as well as higher quality field verification. Thanks to the recent development, there are already many of high-resolution data recently launched, which can be tapped to complement existing sets. And development of crop-cutting techniques worldwide already shows great promise in transparently, accurately and timely collection of agricultural yield data.

We then show that as this improved quality satellite-based risk information could reflect crop production cycles, losses and risks at granular level but covers the whole system of the country and over time, one can use them to understand localized risks and how they co-moves and the overall risk of the country, all of which are necessary for designing and pricing competitive financial contracts. This necessary understanding could then be used to design optimal PPP arrangements that result of the most cost effective contribution to risk management from farmers, market and government. With better risk information, we show that if government can be more targeted and limit their role to just serve as insurer of last resort providing help in the extreme events when it is not cost effective for market to manage and allows the market to work offering variety of competitive financial contracts to transfer risk that have great values to farmers and banks, this could result in reduction in vulnerability of farm households, improve financial deepening and financial stability of rural agricultural bank and most importantly reduce fiscal burden to government. This is really a great improvement from the current program.

This study is far from inclusive, there are obviously many other possible ways one can exploit from high quality risk information, for example (1) *Improving management of other agricultural risk* more importantly price risk, (2) *Crowding in effectiveness of other public programs* by linking government supports to enhance farmers' incentives to participate in crop zoning, large farm program, farmer's risk sharing groups, etc. or by designing risk transfer mechanisms to protect government/BAAC lending agricultural SMEs and (3) *Enhancing sustainable financial deepening* by allowing BAAC to design variety of insured credit and saving programs and allowing them to do risk-based pricing which could result in the BAAC being able to extend credit to those currently rationed out of market due to potentially high credit risk.

Overall, our results point to the three key implications: (1) *Need to investing in agricultural risk information*: This applies well beyond satellite data into better administrative data collection, combining existing data available at various places. But who should invest in data? And there will be needs for research and development to get high-quality data to function. (2) *Need to investing in research and development*: We already show that great data would still require high degree of research. But who should do research? And even great research outcomes will not matter if it could not be brought into implementation. And (3) *Need to answer "Who should invest in data, who should do research, and how do we bring research to materialize and drive policy?"*

These are key questions; of which we would like to open three debatable ways forward.

Focus on getting the market to work first: Government should limit their role in agriculture to rather support development of healthy market. Key roles of government should be provisioning of public goods, e.g., convening agricultural data reform, investing in risk and market infrastructure (including inviting laws) to incentivize private sector to innovate and compete in the market. This could then later create value of risk information beyond public good, where private sector will have incentive to co-invest both in data and research and development

Reform agricultural data in the country: With government having convening power and with many agricultural data currently collected by various government agencies, government should involve private sectors, researchers at the start in the reforming (and opening) of agricultural data in the country, which involve not only collating existing data but designing data platform and collection going forwards. Inclusive involvement will create sense of ownership and incentivize them to partner in data usage and investment going forward.

Bridging research to market and policy environment: We already show that agricultural policies could work well when optimal public-private partnership (PPP) is in place. So bridging research in, improve research capacity of government officials and more importantly involving both researchers and private sectors at the start of the policy design would be critical.

Reference

Barrett, C.B., B.J. Barnett, M.R. Carter, S. Chantarat, J.W. Hansen, A.G. Mude, D.E. Osgood, J.R Skees, C.G. Turvey and M.N. Ward. 2007. Poverty Traps and Climate Risk: Limitations and Opportunities of Index-Based Risk Financing, IRI Technical report 07-02, IRI, Columbia University.

Binswanger-Mkhize, H.P. 2012. Is there too much hype about index-based agricultural insurance? Journal of Development Studies, 48(2), 187-200.

Carter, M., A. de Janvry, E. Sadoulet and A. Sarris. 2014. Index-based weather insurance for developing countries: A review of evidence and a set of propositions for up-scaling (No. P111).

Carter, M. R. 2013. Sharing the risk and the uncertainty: Public-private reinsurance partnerships for viable agricultural insurance markets. I4 Index Insurance Innovation Initiative Brief.

Chantarat, S. 2015. Index-based Risk Financing and Development of Disaster Insurance Program in Developing Countries. PIER Discussion Paper 10.

Chantarat, S., A.G. Mude, C.B. Barrett and M.R. Carter. 2013. Designing Index-Based Livestock Insurance for Managing Asset Risk in Northern Kenya, Journal of Risk and Insurance 80(1): 205-237.

Chantarat, S., C.B. Barrett, A.G. Mude, and C.G. Turvey. 2007. Using Weather Index Insurance to Improve Drought Response for Famine Prevention, American Journal of Agricultural Economics, 89 (5): 1262-68.

Clarke, D.J., O. Mahul and N. Verma 2012. Index Based Crop Insurance Product Design and Ratemaking. World Bank Policy Working Paper 5986.

Cummins, J.D. and O. Mahul. 2009. Catastrophe Risk Financing in Developing Countries: Principles for Public Intervention, The World Bank.

Froot K.A., ed. 1999. The Financing of Catastrophic Risk. The University of Chicago Press, Chicago and London.

Ghesquiere, F. and O. Mahul. 2010. Financial Protection of the State against Natural Disasters. World Bank Policy Research Working Paper 5429.

Gurenko, E., R. Lester, O. Mahul and S. Oguz Gonulal 2006. Earthquake insurance in Turkey: History of the Turkish Catastrophe Insurance Pool. World Bank Publication.

Guy Carpenter and Company, LLC. 2010. Reinsurance Market Review 2010. Guy Carpenter and Company, New York.

Hofman, D. and P. Brukoff. 2006. Insuring Public Finances Against Natural Disasters: A Survey of Options and Recent Initiatives. IMF Working Paper 06/199.

Ibarra, H. and O. Mahul. 2004. Self-insurance Funds as Agriculture Insurance Providers: The Case of Fondos in Mexico. World Bank mimeo.

IFAD and WFP. 2010. Potential for Scale and Sustainability in Weather Index Insurance, Rome. 33

Karlan, D., R. Osei, I. Osei-Akoto and C. Udry. 2014. Agricultural decisions after relaxing credit and risk constraints. Quarterly Journal of Economics, 129(2).

Khobkhun, B., Prayote, A., Rakwatin, P. and N. Dejdumrong (2013). Rice Phenology Monitoring of Thailand using Time Series MODIS Imagery.

Kunreuther, H. and M. Pauly 2004. Neglecting Disaster: Why Don't People Insure Against Large Losses? Journal of Risk and Uncertainty 28: 5-21.

Mahul, O., and C.J. Stutley. 2010. Government support to agricultural insurance: challenges and options for developing countries. World Bank: World Bank Publications.

Mahul, O. and J.R. Skees. 2007. Managing Agricultural Risk at the Country Level: The Case of Index- based Livestock Insurance in Mongolia. Policy Research Working Paper 4325, The World Bank.

Nelson et al. 2014. Towards and Operational SAR-based rice monitoring system in Asia: Example from 13 demonstration sites across Asia in the RIICE project. Remote sensing 6: 10773-812.

Rakwatin, P., T. Sansena, N. Marjang and A. Rungsipanich. 2012. Using Multi-temporal Remote Sensing Data to Estimate 2011 Flood Area and Volume Over Chao Phraya River Basin, Thailand. Working Paper. GISTDA, Thailand.

Siamwalla, A. and W. Na Ranong. 1980. Rice Handbook, published manuscript in Thai.