Downscaling Seasonal Climate Forecasts for Risks Management of Rice Production in the Philippines

Virgilio Julius P. Manzano Jr.1,* and Amor V.M. Ines2,3,4

1Department of Agricultural and Biosystems Engineering, Mariano Marcos State University, City of Batac 2906, Ilocos Norte, Philippines
2Department of Plant, Soil, and Microbial Sciences, Michigan State University, East Lansing, MI, 48824, USA
3Department of Biosystems and Agricultural Engineering, Michigan State University, East Lansing, MI, 48824, USA
4International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, NY, USA 10964

Abstract

Objectives: This study is centered on the potential use of a dynamic seasonal climate forecast for informing climate risk management in Central Luzon, Philippines to improve rice productivity and resilience. Specifically, we seek to test the downscalability of the seasonal climate forecasts in the region using a multi-variate spatio-temporal downscaling technique, understand and assess the predictability of rice yield at selected growing areas in the Philippines, and provide guidance on how to develop agricultural risk management strategies.

Methods/statistical analysis: The coupled Global Circulation Model (GCM) CFSv2 was used to evaluate the utility of MJJA (May-June-July-August) rainfall forecasts for risk management of rice production in Central Luzon, Philippines. We used a non-homogeneous hidden Markov model (NHMM) to downscale and simulate the GCM forecasts to selected weather stations in the region. On the other hand, we evaluated the skill of the climate forecasts for predicting crop yields. The simulated rainfall was used to drive the rice models set up in DSSATv4.5. Other weather variables needed by DSSAT were generated and conditioned on the occurrence of rainfall based on NHMM rainfall simulation. Simulated rice yields obtained from these models using observed (i.e., simulated by observed weather) and conditioned rainfall (i.e., NHMM downscaling) serve as a baseline for evaluating yields. We also performed a sensitivity analysis to assess appropriate planting windows for the target season for risk management.

Findings: Inter-annual variability of rainfall is moderately simulated, with a skill ($r$) of 0.41, suggesting that NHMM was fairly successful downscaling rainfall from the regional scale given the predictive nature of the predictor, at three months lead-
time. The observed MJJA mean rainfall from the six stations falls at around 56% within the interquartile ranges of simulated rainfall, which indicates reasonable skill of the NHMM downscaling CFSv2’s MJJA rainfall hindcasts. Simulated yield was found comparable with the observed at 5% level of significance using $t$-test ($p > 0.05$). The correlations between observed and predicted yields are equal to 0.56. This indicates that the models can represent about 31.36% of the inter-annual variability of the yields of rice, albeit of the three months lead-time of the CFSv2 hindcasts. It suggests a reasonable performance of the models in simulating rice yield using the NHMM generated climate information. Climatologically, the best planting and sowing windows for rice in the study area is on the first week of May. This can be adjusted by using seasonal climate forecasts information. Harvest period should not cross over in the month of September to avoid exposure to heavy typhoons. **Application/improvements:** Sensitivity analysis showed that planting rice earlier than the usual planting windows practiced by the farmers could improve resilience to climate risks. Managing the variance of this management window, however, is of paramount importance, which can be informed by skillful climate forecasts.

**Keywords:** Seasonal Climate Forecast, Downscaling, Non-homogenous Hidden Markov Model, Crop Model, Climate Risks Management, Yield Prediction.

### 1. Introduction

The persistent exposure of the Philippines to climate extremes contributes to its developmental problems. Agriculture, being the primary development sector contributing more than 20% to the gross domestic product (GDP), is particularly sensitive to climate. The impacts of climate-related hazards, such as droughts and floods, have enormous social and economic consequences at the farmer’s level and to the national economy.

Climate variability and extremes are very much a part of life in the Philippines. Where people are poor and vulnerable, as in farming communities, these climatic factors add to the challenges of life. If crops fail due to climatic shocks, smallholder farmers have few or no alternative means of providing the basic needs for their families. Climate presents risks at the individual, and national level. However, it also presents opportunities that can be exploited [1].

Managing impacts of climate variability and extremes in agriculture is of utmost importance not only to minimize the risk associated with the bad years, but also to maximize opportunities during the good years. The bad years potentially threatens established aspects of farming systems and may require adjustments to current practices in order to maintain productivity. Many studies focus on the negative impacts of extreme events (e.g., super-typhoons). Nonetheless, it can also bring benefits. In semi-arid regions of the Philippines, such as in Northern Luzon, a large portion of the annual rain comes...
from cyclones. Cyclones provide relief from dry-spells, replenishes the groundwater, and prevents potential salt water intrusion.

In order to manage better climate-related risks in agriculture, one must be equipped with useful information in advance to plan ahead in time. Advances in climate science using advanced earth system models combined with statistical models have paved the way to predict climate in advance of the growing season [2–3]. This advanced climate information could help farmers make better decisions to minimize climate risks and exploit climate opportunities, although it poses a lot of challenges [4].

Forecasting, however, does not always guarantee a desired outcome. It encompasses predictable variability as well as inherent uncertainty. For example, the rainy season is a predictable occurrence, but the amount, timing, and distribution of the rains is uncertain. Nonetheless, advances in climate science are improving the predictability of climate variability, which is beneficial to protecting societies against unfavorable climate events. Weather forecasts have had major advances in helping plan appropriately at the shorter time scales. Moreover, seasonal forecasts are potentially very useful for planning agricultural activities and as a starting point for early warning and response planning [5].

Much of the theoretical basis of seasonal climate predictions is predicated on predictability of large-scale tropical sea surface temperature (SST) anomalies at seasonal lead times. Climate forecasts are very much associated in varying tropical ocean conditions; measurements of SST in the tropical pacific ocean are especially useful. By measuring these conditions, it is possible to predict climate variability up to several months into the future [6–7].

The El Niño Southern Oscillation (ENSO), which is in the form of El Niño (warmer temperatures) or La Niña (cooler ones), represents one of the most important sources of predictability for climate forecasting, whose effects are very predictable, which on the average, happen every three to seven years. Fluctuations in temperatures can create large changes in evaporation of the sea surface. ENSO event then starts and propagates a process that affects rainfall patterns in the particular region. ENSO is the most significant source of seasonal climate variability globally [3]. The rainfall over the Philippines is strongly influenced by ENSO [8].

Limited work has been done regarding the usage of seasonal climate forecasts and crop models over the Philippines, particularly in northern and central Luzon [9–10]. Crop models can simulate the responses of the crops to climate, environmental conditions, and management practices [11]. There is a significant opportunity to reduce production risks inherent to rainfed rice growing by means of the timely disseminations of seasonal climate forecasts and crop yield predictions to farmers as well as to local government officers. When reliable climate forecasts are available, farmers may appropriately decide when, how, and where to plant accordingly. They could also use this information to decide on the type of management practices to be applied. This, in turn, is expected to lower the production risks imposed by climate variability and could result in increased agricultural outputs.

For the 100 million people in the Philippines, rice provides more than half of their caloric intake. Rice is very vulnerable to climate, and in particular to the risk of flooding, especially during the end of the growing season. The annual occurrence of flash floods
and typhoons results in heavy production losses in paddy rice. In the Philippines, farmers experience high risk of crop losses (40–80%) during the typhoon months of June to December (https://www.agriskmanagementforum.org/blog/flood-tolerant-rice-managing-production-risk-philippines). Drought can devastate rice production as well.

2. Materials and Methods

2.1. Study Area

Central Luzon (Figure 1) is a large producer of rice in the country. Here we highlight Nueva Ecija only as they the main producer of rice in the region.

Nueva Ecija is a landlocked province in central Luzon made up of low lying alluvial plains in the west, central, and southwestern parts and rolling uplands in the northern, eastern, and southeastern parts. Nueva Ecija is a first-class province, earning PhP 450 million ($9 million) or more annually, made up of five cities and 27 municipalities. The primary product of Nueva Ecija is rice. In 2007, it produced around 942,613 metric tons of milled rice (http://countrystat.bas.gov.ph) while in 2014 it produced 953,292 metric tons. Temperatures in Nueva Ecija range from a minimum average of 22 °C during the cooler months of November to March to a maximum average of about 32 °C, occurring between April and July. Three climatic types are found in the province; Type I, with two

![Figure 1](image_url)
pronounced seasons, wet and dry; Type III, with seasons not very pronounced; and Type IV, with rainfall evenly distributed throughout the year (www.pagasa.dost.gov.ph).

2.2. Downscaling Climate Forecasts and Crop Simulations

Figure 2 shows the general framework followed in this study. Climate forecasts can come from SST-based approach or by using directly GCM rainfall as predictor of the forecast. Here, we present the use of a dynamic seasonal climate forecast for informing agricultural risk management at the policy or farmer level. A crop model was used to ingest the downscaled seasonal climate forecasts to predict yields.

The stochastic simulations of daily rainfall were used as inputs to the crop models, simulating the crop growth for each of the 100 NHMM simulations, over the period 1982–2013. Specifically, we focused on Nueva Ecija (station 6). Daily maximum and minimum temperatures and solar radiation are set at their monthly climatological means, conditioned on the occurrence or non-occurrence of rainfall. We used the properties of the dominant soil type and crop cultivars in the study sites. The resulting observed yields are then compared with observed yields simulated with observed weather.

2.2.1. GCM Forecasts

Rainfall forecasts (here, hindcasts) for the period May–June–July–August (MJJA) were derived from the “NOAA-National Centers for Environmental Protection” Climate Forecast System Version 2 (CFSv2 – coupled GCM) using a 3-month lead-time, i.e., derived from predictions initialized in February 1 [12]. The selected model domain encompassed latitudes 14N–20N and longitudes 118E–124E (56-grids) for the period 1982–2013. The extracted seasonal time series is comprised of the mean of 24 CFSv2 ensemble members within this domain.

![FIGURE 2. General framework of the study.](image-url)
2.2.2. Non-homogenous Hidden Markov Model

We used the Non-homogenous Hidden Markov Model (NHMM; Figure 3) to downscale predicted CGM-CFSv2 rainfall in space and time. NHMM is a multivariate, stochastic simulation model that simulates stochastic local-scale daily time series (\( R \)) based on a small set of “hidden” states (\( S \)) defined from daily rainfall observations at a network of stations [13–15]. It factorizes the joint distribution of historical daily rainfall amounts by making two assumptions: first, that the rainfall on a given day only depends on the state active on that day, and second, that the state active on a given day depends only on the previous day’s state. The latter assumption corresponds to the Markov property, while the fact that the states themselves are not directly observable accounts for the “hidden” in the model description. Downscaling is achieved by allowing the Markovian transition probabilities between the states to vary non-homogeneously over time, according to a set of predictors (\( X \)). Once the NHMM’s parameters have been learned, stochastic simulations of rainfall can be generated at all the stations on the network [16]. We assumed a conditional independence (C.I.) between or among stations (\( R \)) when we generated rainfall amounts for each station once a rainfall event is simulated. The selected rainfall stations in the region are not too far from each other (Figure 1).

2.2.3. Observed Climate Data

To calibrate the NHMM, 32-year (1982–2013) daily rainfall amounts (from May 1st to August 31st) from six PAGASA (Philippine Atmospheric Geophysical and Astronomical Services Administration) agro-meteorology and synoptic weather stations in northern Philippines were used (Figure 1). A difference in the temporal distributions of rainfall is apparent among the selected stations (Figure 4a). The Batac, Laoag, Baguio, and Nueva Ecija stations have two pronounced seasons: dry from November to April and wet during the rest of the year. Isabela and Tuguegarao stations are more or less with evenly distributed rainfall, with a relatively dry climate from January to April. Generally, rainfall frequency

![Figure 3. Schematic of NHMM (adopted from Kirshner, 2005; Robertson et al., 2007).](image_url)
FIGURE 4.  (a) Monthly rainfall amounts of each station (n = 32 years). (b) The mean seasonal cycle of rainfall frequency (days/decade) computed using decade block means. (c) The mean seasonal cycle of rainfall intensity (mm/day) at each station, computed using decade block means.
increases from May, reaches a peak in July and recedes in August (Figure 4b). As shown in Figure 4c, rainfall intensity is fairly uniform for the Nueva Ecija, Isabela, and Tuguegarao stations, while in Baguio, Batac, and Laoag stations it tends to increase from May, reaching a peak in August. The differences in climate are attributed to the presence of a central mountain range and their exposure to the southwest and northeast monsoons.

### 2.2.4. GCM Data Bias Correction and Modulation

Monthly CFSv2 rainfall downloaded from the International Research Institute for Climate and Society – IRI Library (www.iridl.ldeo.columbia.edu) was bias corrected using a quantile-to-quantile mapping technique [17–19]. Then, the bias-corrected monthly GCM rainfall was converted to daily intensities (Figure 5a). This time series was used as a predictor of the NHMM. When the uniform rainfall intensity was used as a predictor, NHMM runs showed that the simulated seasonal rainfall tends to follow a flat pattern (not shown). Because of this, we applied a modulation technique to the monthly uniform rainfall intensities as shown in Figure 5b (sample, 1982); monthly GCM rainfall was multiplied with a modulation daily factor estimated from observed climatology. Modulation was necessary to closely mimic the temporal pattern of the observed station rainfall.

### 2.2.5. Training the NHMM

The rainfall model parameters for each station, number of large-scale hidden “weather” states, and transition probability matrix of these “weather” states, are the major NHMM parameters that need to be calibrated before downscaling. The rainfall model for each

![FIGURE 5](image-url) **FIGURE 5.** GCM hindcasts, lead time 3-months: a) uniform rainfall intensity, MJJA 1982; b) modulated rainfall intensity, MJJA 1982.
station was characterized by a combination of Dirac-delta function, for rainfall occurrence, and hyper-exponential distribution, for rainfall intensity. The Expectation–Maximization (EM) was used by NHMM to learn and estimate those statistical model parameters using the patterns of daily rainfall data (1982–2013) from the selected weather stations (Figure 1). First, we assumed a number of hidden states, and run the calibration. EM maximized the likelihood function. We repeated this process until we found the optimal number of states, with optimized statistical model parameters; Bayesian Inference Criterion (BIC) was used to measure the performance of the best statistical model for northern and central Luzon, Philippines.

2.2.6. Downscaling and Linking Climate Forecasts with a Crop Model

The calibrated NHMM model was used to downscale large-scale GCM data. But first, we examined the ability of the NHMM to downscale rainfall in space and time. A perfect prognosis analysis was conducted using monthly time series of observed rainfall as a predictor. The monthly rainfall time series were averaged from the selected stations in the region, and then modulated to daily intensities before we used them as a predictor. We measured NHMM performance based on downscaled daily rainfall in each station, and across the region. Then, we conducted the downscaling of actual GCM data using the predicted CFSv2 MJJA seasonal rainfall in northern and central Luzon (1982–2013).

The simulated rainfall was used to drive the rice models set up in DSSATv4.5 [20–21]. Other weather variables needed by DSSAT were generated and conditioned on the occurrence of rainfall based on NHMM rainfall simulation. Simulated rice yields obtained from these models using observed (i.e., simulated by observed weather) and conditioned rainfall (i.e., NHMM downscaling) serve as a baseline for evaluating yields. We also performed a sensitivity analysis to assess appropriate planting windows for the target season for risk management.

2.3. Analysis of Results

We used Pearson's correlation ($r$) to quantify the performance of NHMM in replicating monthly and seasonal rainfall in northern and central Luzon. $T$-test was also used to evaluate the difference between observed and simulated rainfall and yields at 5% level of significance.

Simulated yields using observed weather from 1982 to 2013 were correlated with simulated annual yields using NHMM downscaling of GCM forecasts.

3. Results and Discussions

3.1. NHMM Hidden Weather States

Figure 6 shows the BIC values optimized by the EM algorithm for every tested NHMM model with corresponding number of hidden states. A six-hidden-state NHMM best described the climate and weather patterns in northern and central Luzon.
The patterns of rainfall intensity (mm/wet-day) and rainfall occurrence (wet-day/day), of each station, for each hidden state, are shown in Figure 7a and b. State 6 is a “dry” state with very small rainfall probabilities. State 5 is a very “wet” state, which implies the occurrences of big storms, uniformly at all stations. This state may be associated with typhoons in the region. States 1 and 4 are characterized by mean rainfall intensities, 10 mm to 20 mm/day, and appear to be like complimentary states. States 2 and 3 exhibit “semi-wet” states with rainfall intensities <10 mm/day. When state 3 occurs, its occurrence of rainfall is slightly more than that of state 2.

A most-likely state sequence for a six-hidden-state NHMM was generated using Viterbi algorithm (Viterbi, 1967). These sequences of rainfall states provide a synoptic view of large-scale weather conditions across the study region on a daily basis (Figure 8a). It allows the interpretation of the 32-yr rainfall record by assigning to each day the state that was most probable on that day, based on a six-hidden-state Hidden Markov Model. The average normal frequencies of the six states are 22.5%, 23.5%, 15.2%, 16.4%, 9.0%, and 13.5%, respectively. The state sequence suggests a strong seasonality, with the very dry state (state 6) dominating on the first week of May and continuous to decline as the wet season progresses. The month of May is where the onset of wet season occurs. June is dominated by states 2 and 3 (semi-wet) while July is dominated by a normal and fairly wet states (1 and 4). State 5 is the wettest state that occurs mostly in the month of August, the peak of the typhoon season (Figure 8b).

### 3.2. NHMM Rainfall Simulations

#### 3.2.1. Diagnostic Runs

Once the statistical model parameters have been estimated, the model was used to simulate rainfall. The ability of NHMM to simulate and downscale rainfall was diagnosed first by downscaling observed regional rainfall (1982–2013) to the selected stations. Here, we set 100 realizations, which produced 100 datasets of 32 sequences of 123 days. NHMM simulated MJJA rainfall was comparable with observed MJJA regional rainfall ($r = 0.93$;}

![BIC of NHMM for northern and central Luzon as a function of number of states](image-url)
Mean Rainfall Intensity, mm/day

Rainfall Occurrence, wet-day/day

FIGURE 7. (a) Hidden state’s rainfall intensity patterns at each station. (b) Hidden state’s wet day probability patterns at each station.

not shown) suggesting that the model is working reasonably well for the region. A scatter plot of July rainfall from the six weather stations and NHMM simulations is shown in Figure 9.
3.2.2. Downscaling GCM Hindcasts

The box-plot of NHMM rainfall made from downscaling CFSv2 MJJA hindcasts is shown in Figure 10. Inter-annual variability of rainfall is moderately simulated, with a skill ($r$) of 0.41, suggesting that NHMM was fairly successful downscaling rainfall from the regional scale given the predictive nature of the predictor, at three months lead-time.

The observed MJJA mean rainfall from the six stations falls at around 56% within the interquartile ranges of simulated rainfall, which indicates reasonable skill of the NHMM downscaling CFSv2’s MJJA rainfall hindcasts (Figure 10).

Summary of monthly and season correlations between observed and NHMM simulated rainfall. As mentioned above, the overall correlation for the region for MJJA season is...
The target areas, Nueva Ecija (station 3) and Isabela (station 4), exhibited adequate correlation of 0.43 and 0.49, respectively. At monthly level, t-test indicated significant correlations only in July at three locations, Batac (station 2), Nueva Ecija, and Isabela. This is true also for the MJJA season, with relatively increased skills ($r = 0.46, 0.43, 0.49$, respectively). Regional predictive skill ($r = 0.41$) is statistically significant. These values indicate a reasonable basis for its utility to drive crop models such as the DSSATv4.5 for risk management in agriculture.

### FIGURE 9.
Sample scatter plot of observed and simulated July rainfall from six weather stations and NHMM simulations.

### FIGURE 10.
Observed (average from six stations) and NHMM simulated MJJA rainfall (box-plots, 100 realizations).
3.3. Yield Simulations

3.3.1. Downscaled Yields

We used DSSATv4.5 [20] to generate simulated yields of rice (with CSM-Rice) and maize (with CSM-Maize) using the climate information produced by NHMM. For comparison, we used yields simulated using observed weather as “observed yields”.

Figure 11 shows the box-plot of the simulated station level rice yield from 100 simulations. Simulated yield was found comparable with the observed at 5% level of significance using t-test ($p > 0.05$). The correlations between observed and predicted yields are equal to 0.56. This indicates that the models can represent about 31.36% of the inter-annual variability of the yields of rice, albeit of the three months lead-time of the CFSv2 hindcasts. It suggests a reasonable performance of the models in simulating rice yield using the NHMM generated climate information. Meanwhile, lower yields were associated with warm ENSO events. Seasonal predictions of precipitation made with general circulation models (GCMs) are often skillful for some regions and seasons, particularly during El Niño–Southern Oscillation (ENSO) events [3], and Philippines’ climate has a strong teleconnection with ENSO [8,22]. The increased correlations in predicted yields as compared with the predicted rainfall could be attributed to the non-linearity of the soil–plant–weather relationships, suggesting further the usability of dynamic crop models for yield predictions and climate risk management.

3.3.2. Analysis of Planting Windows

We examined the best sowing window for rice in the study area. Focusing on Nueva Ecija (station 6), we run DSSATv45 for 32 years using observed weather as inputs to the crop models. Figure 12 shows the box-plot of yields of rice as a function of sowing windows. Depending on the risk-aversion level of the farmer, for example, based on climatology, the best planting windows for rice is on the first week, and on the last week of May, respectively. Seasonal climate forecasts could inform this decision for an upcoming season. Based on

![Figure 11. Observed yields (average from six stations) and yields from NHMM driven DSSAT rice model at Nueva Ecija (station 6; Figure 1).](image-url)
the climate forecast, farmers could adjust planting windows or even decide on which crop variety or varieties to plant for the growing season to minimize risk.

4. Summary and Conclusions

We have used a nonhomogeneous hidden Markov model (NHMM) in conjunction with a crop model to investigate spatial and temporal disaggregation of seasonal rainfall for simulating rice yield in Central Luzon, Philippines. The CFSv2 MJJA rainfall with 3-months lead-time, averaged over 56 grids across 32 years of record was used as the predictor of the NHMM to investigate its ability to downscale rainfall forecasts to station-scale conditions. Stochastic simulations reveal that NHMM is able to recover the inter-annual variability of station scale rainfall modestly ($r = 0.41$). This indicates a reasonable downsampling of CFSv2 regional-scale rainfall to the station scale, given the predictive nature of the predictor data, as well as the imperfect capability of the NHMM. Diagnostics of the NHMM show a skill, $r = 0.93$, which may be attributed to unpredictable station scale noise.

With the reasonable skill of the NHMM in the study region established, the downscaled rainfall simulations were then used to drive a crop model (CSM-Rice) in the DSSATv4.5, to evaluate the performance of the NHMM’s rainfall simulations in terms of simulating rice yield. Simulated yields were found to be moderately correlated with observed yields with $r = 0.56$. Moreover, simulated yields were found comparable with observed at 5% level of significance using t-test ($p > 0.05$). In principle, the results of this study suggest that regional rice yields over the study area could be predicted by CFSv2 rainfall at 3-month lead-time.

**FIGURE 12.** Rice yields as affected by sowing dates simulated by DSSAT-CSM-Rice in Nueva Ecija (station 6).
The main goal of this study is to provide seasonal yield forecast for the upcoming wet growing season for rice in Nueva Ecija, Philippines. Climatologically, the best planting and sowing windows for rice in the study area is on the first week of May. This can be adjusted by using seasonal climate forecasts information. Harvest period should not cross over in the month of September to avoid exposure to heavy typhoons.

Acknowledgements

We acknowledge the Ilocos Norte Science Community (INSC) partner agency coordinators for collecting climate data used in this study; PAGASA for providing those datasets; Fulbright-Philippines and the Fulbright program for the research grant provided to VJPM; IRI/Columbia University for hosting VJPM; and special thanks to Bradfield Lyon, Paula Gonzalez and Andrew Robertson for their technical guidance during conduct of study. Link to NHMM Toolbox used in this study: http://iri.columbia.edu/our-expertise/climate/tools/hidden-markov-model-tool/hmmtool/.

References


