Applicability of meteorological ensemble forecasting to predict summer cold damage in rice growth

Ryuhei YOSHIDA\textsuperscript{a,b}, Shin FUKUI\textsuperscript{b,c}, Shin FUKUI\textsuperscript{d,e} and Takeshi YAMAZAKI\textsuperscript{d}

\textsuperscript{a}Faculty of Symbiotic Systems Science, Fukushima University, Fukushima, 960–1296, Japan
\textsuperscript{b}Research Center for Fisheries Resources, National Research Institute of Fisheries Science, Yokohama, 236–8648, Japan
\textsuperscript{c}Faculty of Human Sciences, Waseda University, Tokorozawa, 359–1192, Japan
\textsuperscript{d}Graduate School of Science, Tohoku University, Sendai, 980–8578, Japan
\textsuperscript{e}present affiliation: Japan Meteorological Business Support Center, Tsukuba, 305–0052, Japan

Abstract

Abrupt temperature drops pose serious concerns for rice production in northern Japan. Previous early warning systems have been based on projected temperature tendencies, and alerts have announced for the occurrence of low temperatures. The rice crop has low-temperature-sensitive stages; however, previous systems have not considered them because of the difficulty of simulating rice growth at the local scale. The forecast system would be more valuable by considering both the rice growth stage and current meteorological forecast techniques. In this study, we synthesized ensemble numerical weather prediction and a cultivar-based rice growth model to forecast 14-day cold damage risk. The ensemble mean forecast with nine members predicted surface air temperatures more skillfully for seven days with lower root-mean-square errors (RMSEs) (1.3–1.9°C) than those of the climatological forecast (2.1–2.4°C) that is derived from historical observations over 30 years. The single deterministic forecast predicted the temperatures better for five days with 1.3–2.0°C of RMSEs, showing the extension of the predictable period by two days with ensemble forecasting. For the cooling degree-days, both the single and ensemble mean forecasts showed lower RMSEs than the climatological forecast throughout the forecast period of 14 days (4.1, 3.8, and 5.2°C at the forecast time = 14 day for single deterministic, ensemble mean, and climatological forecasts, respectively). Although the climatological forecast estimated the rice growth stages reasonably, the performance for cooling degree-days was inferior to the ensemble mean and single deterministic forecasts. The meteorological mean state is sufficient to estimate the rice growth stage, but an accurate temporal pattern of the surface air temperature provided by numerical weather forecast is essential for reliable cold damage forecasting. Moreover, ensemble forecasting is more effective than the single deterministic forecast to reduce prediction errors for both the surface air temperature and cold damage.

Key words: Climatological forecast, Ensemble simulation, Low temperature, Predictability, Rice growth simulation

1. Introduction

In Japan, one of the top 10 rice producers worldwide (Ray \textit{et al.}, 2012), rice production is facing extreme high/low temperature risks. High temperatures have recently caused yield decreases and quality deterioration in the southern part of Japan (Hasegawa \textit{et al.}, 2011; Okada \textit{et al.}, 2011), and low temperatures brought by the Yamase flow (cool summer northeasterly flow from the Okhotsk High) have caused yield decreases in the northern part of Japan (Shimono, 2011). It is expected that the cold damage in the present climate would weaken with global warming, and rice production in the current cold area would be more important for an increasing rice yield (Nakagawa \textit{et al.}, 2003; Easterling \textit{et al.}, 2007; Shimono, 2011). However, previous studies reported that the Yamase flow occurs even in the future climate (Kanno \textit{et al.}, 2013; Kanda \textit{et al.}, 2014), and cold sterility under global warming still prevails over heat sterility (Yoshida \textit{et al.}, 2015). Therefore, the numerical forecast of cold damage continues to be an essential factor for stable rice production (Wassmann \textit{et al.}, 2009).

The risks of cold damage depend not only on the meteorological conditions but also on the growth stage of rice plants. Rice at the booting stage is the most sensitive to low temperatures, which cause sterility and yield decreases (Shimono, 2011). However, previous studies have only focused on temperature tendencies to warn of abrupt temperature changes (NARO, 2019; Japan Meteorological Agency [JMA], 2019a). Kobayashi \textit{et al.} (2010) considered local cropped cultivar, but the calculated growth stage was used for the estimation of crop disease risk, not for that of cold damage risk. Moreover, these warning systems in previous studies have provided risks in a qualitative manner.

Difficulty in cultivar-based regional/global simulations hinders us in the information of a detailed rice growth forecast that reflects local rice growth stages; elaborate and laborious field experiments are required for the estimation of rice phenology parameters. To overcome this difficulty of cultivar-based simulation, Fukui \textit{et al.} (2015) calculated the rice phenology parameters for major Japanese cultivars. The sensitivity of daily rice growth to air temperature and daylength
were parameterized for each cultivar. Using these parameters, Yoshida et al. (2015) showed the applicability of a cultivar-based rice growth simulation to assess the impact of climate change.

Modern meteorological forecasts are based on numerical weather prediction, which includes inevitable uncertainties from the initial conditions and forecast model. Ensemble forecasting, which includes multiple integrations from slightly perturbed initial conditions (with different model configurations as well) to represent the uncertainties (e.g., Matsueda and Nakazawa, 2015), is often applied to the currently operational numerical weather forecasts. The mean ensemble forecasts can suppress random errors in the single deterministic forecast. Moreover, the deviation of the ensemble members can predict uncertainties in the forecasts.

Current meteorological forecasts and the cultivar-based simulation enable a rice growth ensemble prediction that reflects the local meteorological conditions and the cropped cultivar. This study aims to evaluate the advantage of applying ensemble numerical weather prediction to a cultivar-based simulation for the estimation of cold damage risk.

2. Materials and Methods

2.1 Meteorological features

We analyzed Tohoku region in northeastern Japan, shown in Fig. 1a. The Ōu mountain range, located in the central part of the region, creates regional meteorological characteristics. In summer, the yamase flow, associated with the Okhotsk high, brings northeasterly winds and cloudy, cool weather to the Pacific coast of northern Japan (Kojima et al., 2006). Because the westward propagation of low-level clouds is blocked by the Ōu mountain range, there are west-east contrasts in climate (Shimada and Kawamura, 2011). The daily maximum temperature in the Pacific Ocean side was, on average, ~2°C lower than that of the Sea of Japan side, according to climatological measurements from 1981 to 2010 (JMA, 2019b). Moreover, the sunshine hours on the Pacific Ocean side are approximately 20–30 h less than on the Sea of Japan side. This west-east contrast due to the cool northeasterly flow is one of the meteorological characteristics in the analysis region.

2.2 Forecast system for summer cold damage

The proposed forecast system in this study was based on numerical weather prediction and a cultivar-based rice growth model. The cooling degree-day on a forecast day was provided from a rice growth simulation with meteorological forecast data. As rice growth stages reflected the previous meteorological history at one local site, the estimated cold damage risk was different among the sites even when the same air temperature was forecasted. Moreover, we processed the forecasts with meteorological ensemble data. Each meteorological member provided different surface air temperature and cooling degree-days. By taking mean value over the members, the performance was expected to improve to that of single member forecast.

2.2.1 Numerical weather prediction

We applied the ensemble numerical weather prediction data with a 5-km grid spacing developed by Fukui et al. (2014). The data are generated by dynamically downscaling each member of the JMA’s global one-month ensemble hindcast data composed of nine members at a 1.25° × 1.25° horizontal resolution with the JMA’s non-hydrostatic model (NHM; Saito et al., 2007) applying the one-way double nesting approach. The outer domain covered by the NHM with a 25-km grid spacing is Japan and its...
surrounding area; the inner domain covered by the NHM with a 5-km grid spacing is Tohoku district, as shown in Fig. 1a. Note that the domains are slightly different from those used by Fukui et al. (2014). The dynamical downscaling forecasts initialized at July 10, July 20, July 31, August 10, and August 20 in 2000–2009 and run for 14 days. See Fukui et al. (2014) for more details.

The simulated meteorological data generally include some biases because of the imperfections in numerical models (e.g., Yoshida et al., 2012). Thus, biases of the ensemble downscaled forecast data should be removed before the analysis. We calculated the bias (b) of the forecasts in each year (yr), initial date (init), forecast time (FT), and site (site) as follows:

\[
b_{yr,init,FT,site} = \frac{1}{N_r - 1} \sum_{r=1}^{N_r} (X_{yr,init,FT,site} - \bar{X}_{yr,init,FT,site}),
\]

where \(N_r\) is the number of analysis years (i.e., 10), \(X\) is the ensemble mean of nine members, \(Y\) is the Mesh-AmeDAS (Automated Meteorological Data Acquisition System) observed value (Seino, 1993), and \(t\) is the training year. Because the independence between corrected and trained data was maintained, we calculated the biases from the data for nine years, excluding the data in the target correction year. This process leads to different biases in the same month and day among the analyzed years. Substituting the difference between the ensemble mean and observations from each ensemble member, we corrected the biases of the ensemble forecast.

### 2.2.2 Rice growth model

We then simulated rice growth with the meteorological ensemble data. The Hasegawa/Horie (H/H) model (Yoshida et al., 2015), which takes into account the difference in rice growth among cultivars, was applied in the simulation. The H/H model first calculates the daily growth rate (development rate; DVR) and then calculates the development index (DVI) by integrating the DVR from the emergence to the j-th day (DVI; 0 for seeding emergence, 1 for panicle initiation, 2 for heading, and 3 for maturity) as follows:

\[
DVR = \sum_{k=0}^{j-1} DVR_k,
\]

where \(k\) is the number of days after emergence, \(T_o\) is the daily mean temperature (°C) and \(L\) is the photoperiod (h). The response functions are as follows:

\[
f_1(T_o) = \begin{cases} \left( \frac{T_o-T_1}{T_2-T_1} \right)^{\alpha}, & (T_o < T_1) \\ 0, & (T_o < T_a \text{ or } T_o > T_a) \end{cases},
\]

\[
g(L) = \begin{cases} \left( \frac{L}{L_o} \right)^{\beta}, & (L_o < L < L_n) \\ 1, & (L < L_o \text{ or } L > L_n) \end{cases},
\]

\[
f_2(T_o) = \begin{cases} \frac{1}{T_o} \left[ 1.0 - \exp \{ -A \times (T_o - T_c) \} \right], & (T_o < T_c) \\ 0, & (T_a \leq T_o) \end{cases},
\]

where \(T_o\) and \(T_a\) are the daily maximum and minimum temperatures for growth thresholds (42°C and 8°C), respectively, and \(L_o\) and \(L_n\) are the same as \(T_o\) and \(T_a\) but for the photoperiods of 24 h and 10 h, respectively. The remaining factors, \(\alpha, \beta, A, G_o, G_a, T_o, T_a, T_c, T_c\), and \(T_a\) are cultivar-dependent parameters: sensitiveness to temperature and photoperiod before heading for \(\alpha\) and \(\beta\); respectively; multiplicative factor of temperature impact on growth after heading for \(A\); minimum number of days required from emergence to heading under optimum conditions for \(G_o\); insensitiveness to temperature after heading for \(G_a\); optimum temperature before heading for \(T_o\); and minimum temperature required for growth for \(T_a\). At each grid, we set the cropped cultivar that had the maximum cropping area in the local prefecture (Fig. 1b) and used the cultivar parameters shown in Table 1.

Although abrupt temperature decreases would be an index of cold damage, both the intensity and period at the booting stage \((1.5 \leq \text{DVI} \leq 2.2)\) are critical for the rice growth (Shimono and Kanno, 2012). Therefore, instead of temperature decrease itself, the cooling degree-days were introduced to consider the cold intensity (Horie et al., 1995). The cooling degree-days \((Q_o)\) at the j-th day is expressed by accumulating cold stress from the threshold \((22^\circ\text{C})\) as follows:

\[
\Delta T_a = \max \left( 22 - T_a, 0 \right), \quad \left( 1.5 \leq \text{DVI} \leq 2.2 \right),
\]

\[
\left( \text{DVI} < 1.5, 2.2 < \text{DVI} \right),
\]

\[
Q_o = \sum_{k=0}^{j-1} \Delta T_a, \quad \left( 0 \leq \text{DVI} \leq 2 \right),
\]

\[
\left( 2 < \text{DVI} \leq 3 \right).
\]

### Table 1. List of cultivars and parameters used in the Hasegawa/Horie (H/H) model. All parameters are the same as those in Yoshida et al. (2015), whereas those of ‘Masshigura’ (MSG) were re-calculated using the method of Fukui et al. (2015).

| Name (Abbreviation) | \(G_1\) (days) | \(G_2\) (days) | \(\alpha\) | \(\beta\) | \(A\) (×10³) | \(T_o\) (°C) | \(T_a\) (°C) |
|---------------------|----------------|----------------|-----------|---------|-------------|------------|------------|
| ‘Masshigura’ (MSG)  | 59.8           | 37.5           | 0.55      | 0.46    | 1.1         | 34.6       | 5.3        |
| ‘Hitomebore’ (HTM)  | 56.7           | 23.3           | 0.93      | 0.95    | 3.5         | 32.9       | 0.4        |
| ‘Akitakomachi’ (AKT) | 54.6       | 24.7           | 1.24      | 0.13    | 5.8         | 34.3       | 7.4        |
| ‘Haenuki’ (HEN)     | 55.1           | 24.0           | 1.68      | 1.12    | 3.7         | 30.0       | 1.5        |
| ‘Koshihikari’ (KSH) | 36.6           | 29.1           | 1.11      | 3.42    | 5.3         | 34.6       | 0.2        |
| ‘Asashinoyume’ (ASH) | 30.9       | 35.8           | 1.07      | 6.70    | 9.9         | 31.3       | 4.3        |

Note: \(G_1\) is the minimum number of days required from emergence to heading under optimum conditions; \(G_2\) is the insensitiveness to temperature after heading; \(\alpha\) and \(\beta\) are the sensitiveness to temperature and photoperiod before heading, respectively; \(A\) is the multiplicative factor of temperature impact on growth after heading; \(T_o\) is the optimum temperature before heading; and \(T_a\) is the minimum temperature required for growth.
The threshold, 22°C, is derived by fitting the calculated cooling degree-days on the observed cold sterility rate (Horie, 1988). The rice growth simulation was processed for a grid with paddy area > 1% in the 10-km grid in 2000 (MLITT, 2012).

2.3 Experimental design

By synthesizing cultivar-based rice simulation and numerical weather ensemble forecast for the estimation of cold damage, we designed an experiment system for summer rice growth (Fig. 2). Taking 14 days, starting at July 11, July 21, August 1, August 11, and August 21 of 2000–2009 (total 140 days) for analysis period, we estimated the daily cold damage. The proposed experiment included the following processes. (1) To set the same initial DVI value among the experiments, the H/H model first input the Mesh-AMeDAS observation data by the day before the forecast experiments. This provided common growth stage among the analyzed forecast experiments. (2) We regarded the simulation that Mesh-AMeDAS observation data continuously inputted the H/H model for 14 days as an observation (OBS) experiment. Although observed cold damage or sterility rate is favorable for validation processes, we used the OBS experiment as a reference because such a database is not currently available on a daily basis. Then, we conducted three types of forecasts; SIN, ENS, and CLM forecasts. The SIN and ENS forecast both applied numerical weather forecast data provided by Fukui et al., (2014). The SIN forecast was based on one ensemble member without any meteorological perturbations, and the CLM forecast was processed from all nine members. The ENS forecast, ensemble forecasts independently performed for each member and outputs were summarized by taking ensemble means. To understand the advantage of numerical weather forecasts on cold stress estimations, it is necessary to compare to the method that does not use numerical weather forecasting. By using historical long-term observation data, the climatological value was derived as a reference. The SIN and ENS forecasts expected to outperform than this climatology-based method. We defined the method as the CLM forecast. Considering a climatological value is generally based on the 30 year average (Arguez and Vose, 2011), we inputted observation data from 1980–2009 to the H/H model and calculated the climatological value by averaging over the 30 years.

2.4 Indices for evaluation of forecast system

To assess the performance of the SIN, ENS, and CLM forecasts, we calculated the root-mean-square error (RMSE) between simulated and observed values, which has been widely used for the evaluation of numerical simulations (e.g., Klein and Lewis, 1970; Simmons and Hollingsworth, 2002; Weisheimer et al., 2011), as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2},$$

Fig. 2. Experimental design. Using Mesh-AMeDAS observation to the Hasegawa/Horie (H/H) rice growth model, the development index (DVI) before the forecast date is prepared to the same value among the four forecast experiments (observation [OBS], single [SIN], ensemble [ENS], and climatology [CLM]). For the periods of forecast experiment, the OBS experiment utilized Mesh-AMeDAS observation continuously. Meteorological forecasts of one member without any perturbation in ensemble members was applied in the SIN forecast. The ENS forecast utilized all nine members, and air temperature and cooling degree-days forecasted by each member was aggregated for the ensemble mean. The CLM forecast used 30 years of historical meteorological conditions (1980–2009). As is the case of the ENS forecast, meteorological data at each year were inputted to the H/H model independently. Temperature and cooling degree-days averaged over the simulated 30 years provided climatological value that could be used as a reference.
where \( x_i \) is the simulated value on the \( i \)-th day given by the one forecast of SIN, ENS, and CLM (note: ensemble mean for the ENS forecast, and 30-year average for the CLM forecast), \( y_i \) is the value simulated by the OBS experiment, and \( N \) is the number of samples. The RMSE takes a smaller value for better simulations, and vice versa. The evaluation based on the RMSE was processed on each forecast day, the whole 14 days of forecast period, and three parts that divided the 14 days; short (FT: 1–5 days), medium (6–10 days), and long ranges (11–14 days). Therefore, the parameter \( N \) depends on analysis; 668 (number of analyzed grid cells) for regional analysis, 14 for evaluation of the whole 14 day forecast at each grid cell, and five or four for the three forecast ranges at each grid cell (five days for short- and middle-range forecasts and four days for a long-range forecast). The RMSEs were calculated for each year and initialized the forecast and then averaged over the 10 analyzed years.

For the ENS forecast, we also calculated the ensemble spread \( (s) \) at a given grid cell as follows:

\[
s = \frac{1}{9} \sum_{m=1}^{9} (v_m - a)^2,
\]

where \( v_m \) is the surface air temperature or cooling degree-days of \( m \)-th ensemble member and \( a \) is the ensemble mean over the nine members. The spread is used to understand that the ensemble system has enough variations by comparing the RMSEs (Fortin et al., 2014). As is the case of RMSEs, the spread was calculated for each grid cells and then they were averaged over the analyzed periods.

To quantify the performance of the SIN or ENS forecasts relative to the CLM forecast, we introduced an improvement ratio, \( r_{\text{predict}} \) as follows:

\[
r_{\text{predict}} = \frac{r_{\text{CLM}} - r_{\text{SIM}}}{r_{\text{CLM}}},
\]

where \( r_{\text{CLM}} \) is the RMSE of the CLM forecast and \( r_{\text{SIM}} \) is the RMSE of either the SIN forecast \( (r_{\text{SIN}}) \) or the ENS forecast \( (r_{\text{ENS}}) \). The value of \( r_{\text{predict}} \) takes a positive value when the SIN or CLM forecast performs better than the CLM forecast, and vice versa. The difference between \( r_{\text{CLM}} \) and \( r_{\text{SIM}} \) was normalized by \( r_{\text{CLM}} \) to represent the effectivity of numerical whether prediction relative to the observation-based forecast.

Each simulation is a 14-day integration, and the initial date is shifted every 10 days (except for 11 days between July 20 and July 31). Therefore, there are four overlapping days (i.e., July 21–24, August 1–3, August 11–14, and August 21–24), which are the 11–14th forecast days for one initial condition and the 1–4th forecast days for the next initial condition (Note that they were 12–14th and 1–3rd forecast days between the July 20 and July 31 initialized forecasts). We introduced another index, \( r_{\text{update}} \) to estimate the impact of updating the ensemble forecast to the latest forecast:

\[
r_{\text{update}} = \frac{r_{11-14} - r_{1-4}}{r_{11-14}},
\]

where \( r_{1-4} \) and \( r_{11-14} \) are the RMSE of the ENS forecast for the 1–4th and the 11–14th forecast days, respectively. The \( r_{\text{update}} \) values are positive/negative when the forecast update is effective/ineffective.

### 3. Results

#### 3.1 Geographical distribution of surface air temperature and cooling degree-days

For the forecast period of 1–14 days, the RMSEs of raw surface air temperature averaged over all case were 2–3°C for the most of the analyzed area and 4–5°C for the northeastern or high elevation area (Fig. 3a). These RMSEs were suppressed by applying bias correction (Fig. 3b), decreasing from the regional average of 2.6°C for the raw data to 2.3°C for the bias-corrected data. An improvement in the RMSEs were also found on the cooling degree-days, showing the regional average of 3.2°C to that of 2.4°C (Fig. 3c–d). In the Sea of Japan side, some areas showed the RMSEs as close to zero even for the raw data, although non-zero RMSEs of the surface air temperature were derived in the same area. This was because that the surface air temperature was higher than the threshold to count the cooling degree-days, and the both simulated and observed cooling degree-days were negligibly small in their values.

![Fig. 3](image.png)

Fig. 3. Geographical distribution of root-mean-square errors (RMSEs) for (a) raw and (b) bias-corrected surface air temperature during the forecast time (FT) of 1–14 days. (c–d) Same as (a–b) but for the cooling degree-days. Gray areas indicate non-paddy areas with ≤ 1% paddy fields in 10-km grid cells.
Using the bias-corrected dataset, we then divided whole period of 14 forecast days into the three time ranges. The ensemble mean forecasts of surface air temperature deteriorated with FT. The RMSEs for the short-range forecast (FT of 1–5 days) were from 1–2°C; they were also 2–3°C for the medium-range (FT of 6–10 days), and 2–4°C for the long-range forecasts (FT of 11–14 days) (Fig. 4a–c). The RMSEs increased by approximately 1°C during the forecast period of 14 days. The surface air temperature on the Sea of Japan side was relatively better simulated than that on the Pacific Ocean side. The ensemble spreads showed similar spatiotemporal variation to that of the RMSEs, increasing with the FT and showing a west–east contrast (Fig. 4d–f). The spread to RMSE ratio, which approaches 1 for the ideal system (Takano 2002), was 0.4–0.6 for the short-range forecast, and 0.6–1.0 over the region for the medium- and long-range forecasts (Fig. 4g–i).

For the cooling degree-days, the RMSEs of the ENS forecasts over the short-range forecast were less than 2°C for most parts of the analyzed region (Fig. 5a). While the RMSEs over the medium- and long-range forecasts increased in the northern part of the domain and in the high elevation areas (see Fig. 1 for topography), there were some low-RMSE areas on the Sea of Japan side throughout the forecast period (Fig. 5b–c). As described in analysis of the FT of 1–14 days, daily mean temperatures in the small RMSE areas were generally higher than 22°C, which caused the areas to be free from low-temperature stresses. The increase in RMSEs found in the Pacific Ocean side...

Fig. 4. Geographical distribution of root-mean-square errors (RMSEs) of surface air temperature simulated by the ensemble (ENS) forecast for the forecast time (FT) of (a) 1–5 days, (b) 6–10 days, and (c) 11–14 days. (d–f) Same as (a–c) but for the ensemble spread (sprd) and (g–i) the ratio of sprd relative to RMSEs. Gray areas indicate non-paddy areas with ≤ 1% paddy fields in 10-km grid cells.

Fig. 5. Geographical distribution of root-mean-square errors (RMSEs) of the cooling degree-days simulated by the ensemble (ENS) forecast for the forecast time (FT) of (a) 1–5 days, (b) 6–10 days, and (c) 11–14 days. (d–f) Same as (a–c) but for the ensemble spread (sprd) and (g–i) the ratio of sprd relative to RMSEs. Gray areas indicate non-paddy areas with ≤ 1% paddy fields in 10-km grid cells.
was caused by cold daily mean temperature that was sufficient to count the cooling degree-days. The areas in which the spread increased corresponded well to the RMSE-increased areas, having 1–4°C for the long-range forecast at high elevations and on the Pacific Ocean side (Fig. 5d–f). The spread to RMSE ratio was 0.2–4.0 for the short-range forecast, which increased to 0.4–0.6 over the medium- and long-range forecasts (Fig. 5g–i).

Performances of the surface air temperature and the cooling degree-days were summarized over the grids where the same cultivar was cropped (Table 2; see Fig. 1b for the geographical distribution of cropped cultivars). As a reference, we also showed the temperature difference from the threshold (ΔTc). For each period, surface air temperature showed similar RMSEs among cultivars, while the RMSEs increased with FT, ranging from 1.6–1.8°C in the short-range forecast to 2.6–2.9°C for the long-range forecast. The ΔTc also showed similar temperature in the RMSEs and had larger values for ‘Massigura’ (MSG) and ‘Hitomebore’ (HTM) compared to those for other three cultivars (‘Akitakomachi’ (AKT), ‘Haenuki’ (HEN), and ‘Koshihikari’ (KSH)). These tendencies in increasing RMSEs with the forecast period were also found for Qc. In the long-range forecast, MSG and HTM showed larger RMSEs (~1°C) than did the other cultivars.

### 3.2 Performance of the forecasts for the case of the most intense cold damage

In 2003, the largest cooling degree-days were estimated over the 10 analyzed years (Table 3). Serious cold damage occurred, and the smallest rice yields of the analyzed period were also found for ‘Hitomebore’ in the RMSEs and had larger values for ‘Massigura’ (MSG) and ‘Hitomebore’ (HTM) compared to those for other three cultivars (‘Akitakomachi’ (AKT), ‘Haenuki’ (HEN), and ‘Koshihikari’ (KSH)). These tendencies in increasing RMSEs with the forecast period were also found for Qc. In the long-range forecast, MSG and HTM showed larger RMSEs (~1°C) than did the other cultivars.

### 3.3 Predictability of surface air temperature and cooling degree-days

We then summarized the performance of each forecast starting from the same initial date (Fig. 7). The SIN and ENS forecasts for surface air temperature generally performed better than the CLM forecast for one week and turned during the second week. On average, the effective day that the SIN and ENS forecasts provided better performance than the CLM forecast was five days for the SIN forecast (RMSEs: 1.3–2.0°C) and seven days for the ENS forecast (1.3–1.9°C), while the CLM forecast performed with 2.1–2.4°C of RMSEs during the seven

---

**Table 2.** Summary of root-mean-square errors (RMSEs) of the ensemble (ENS) forecast in simulating surface air temperature (Tc), temperature difference from the threshold of 22°C when the rice is in the booting stage (ΔTc) and cooling degree-days (Qc) aggregated with cultivars (°C).

| FT=1–5 days | MSG | HTM | AKT | HEN | KSH |
|-------------|-----|-----|-----|-----|-----|
| Tc          | 3.6 | 2.1 | 2.1 | 2.1 | 2.1 |
| ΔTc         | 1.8 | 1.0 | 1.0 | 1.0 | 1.0 |
| Qc          | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |

| FT=6–10 days | MSG | HTM | AKT | HEN | KSH |
|--------------|-----|-----|-----|-----|-----|
| Tc           | 2.3 | 1.4 | 1.4 | 1.4 | 1.4 |
| ΔTc          | 1.5 | 1.0 | 1.0 | 1.0 | 1.0 |
| Qc           | 2.3 | 1.0 | 1.0 | 1.0 | 1.0 |

| FT=11–14 days | MSG | HTM | AKT | HEN | KSH |
|---------------|-----|-----|-----|-----|-----|
| Tc            | 2.8 | 2.8 | 2.8 | 2.8 | 2.8 |
| ΔTc           | 1.8 | 1.8 | 1.8 | 1.8 | 1.8 |
| Qc            | 2.6 | 2.6 | 2.6 | 2.6 | 2.6 |

Note: FT, forecast time; MSG, ‘Massigura’; HTM, ‘Hitomebore’; AKT, ‘Akitakomachi’; HEN, ‘Haenuki’; and KSH, ‘Koshihikari’. See Fig. 1b for geographical location of the cropped area for each cultivar (ASH is excluded because of its small geographical share).

**Table 3.** Cooling degree-days (CDD) on August 31 simulated with the Mesh-AMeDAS observations.

| Year | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
|------|------|------|------|------|------|------|------|------|------|------|
| CDD  | 4.0  | 26.4 | 21.6 | 27.1 | 16.7 | 3.6  | 3.7  | 8.8  | 25.2 | 14.3 |
The ensemble spread showed a similar trend among the forecasts initialized at different dates (Fig. 8b). The spread to RMSE ratio increased with the forecast period for all experiment (Fig. 8c), with ratios of 0.75, 0.81, 1.00, 0.91, and 0.78 at the FT of 14 days initialized on the dates of July 10, July 20, July 31, August 10, and August 20, respectively.

Simulated cooling degree-days gradually increased RMSEs with FT and showed higher RMSEs in August than in July (Fig. 8d). The RMSEs were negligibly small throughout the forecast period initialized on July 10 because of the aforementioned insufficient DVI condition (i.e., DVI < 1.5). While the time evolution of RMSEs in the ENS and CLM forecasts were similar, an advantage of the ENS forecast was apparent, with smaller RMSEs throughout the forecast period. As is the case with surface air temperature, the ensemble spreads for the cooling degree-days had a similar tendency to RMSEs, gradually increasing with FT (Fig. 8e). The spread to RMSE ratio for cooling degree-days also increased with the forecast period, showing 0.35, 0.48, 0.71, 0.46, and 0.36 for the initial dates of July 10, July 20, July 31, August 10, and August 20 at the FT of 14 days (Fig. 8f).

3.4 Impact of forecast updating for cold damage

The simulated cooling degree-days in the overlapping days of July 21–24 showed similar RMSEs between the 1–4 days forecast (initial: July 20) and the 11–14 days forecast (initial: July 10) (Fig. 9), showing 0.07°C and 0.11°C, respectively. The difference was apparent in other cases. The RMSEs of the 1–4 days forecasts were approximately 1°C for August 1–3, 11–14, and 21–24, and those of the 11–14 days forecasts ranged from 2.5°C in August 1–3 to 6.6°C in August 21–24. The 1–4 days forecast provided smaller RMSEs than the 11–14 days forecast in all cases. An improvement ratio of the 1–4 days forecast relative to the 11–14 days forecast, \( r_{\text{update}} \) (Eq. 12) had higher value in the forecasts for August (65.3–82.8%) than that for July (36.4%).

4. Discussion

4.1 Advantages of numerical weather forecast to the CLM forecast

The forecast based on numerical weather predictions performed reasonably in simulating cooling degree-days with lower RMSEs than the CLM forecast that utilized historical Mesh-AmEeDAS observations. To understand differences in the performances, we focused on variations in surface air temperature. The amplitude of daily variation of surface air temperature provided by the SIN and ENS forecasts showed better performance than the CLM forecast in simulating the cooling degree-days. Moreover, as seen in the cultivar-aggregated statistics (Table 2), cultivars MSG and
Fig. 7. Time evolution of root-mean-square errors (RMSEs) for the climatology (CLM) (thick line), single (SIN) (dashed line), and ensemble (ENS) (thin line) forecasts of surface air temperature (left) and cooling degree-days (right) aggregated over the same initial dates. The dates shown in the top left of each panel are the initial dates of each forecast.

Fig. 8. Time evolution of (a) root-mean-square errors (RMSEs) of the ensemble (ENS) (solid line) and climatology (CLM) (dashed line) forecasts, (b) ensemble spread among nine members of the ENS forecast, and (c) spread to RMSEs ratio for the simulation of surface air temperature. The color corresponds to the initial dates of the forecasts: purple for July 10, blue for July 20, green for July 31, orange for August 10, and red for August 20. Panels of (d–f) are the same as those of (a–c) but for the cooling degree-days.
HTM showed larger RMSEs for the cooling degree-days than other cultivars, while they had similar RMSEs for the surface air temperature. Averaging over the same cropped area, the observed temperatures during the analysis period were 21.4°C for MSG, 22.0°C for HTM, 22.8°C for AKT, 22.7°C for HEN, and 23.6°C for KSH. The performance deteriorated in the relatively cold region. Considering that the cooling degree-days increases when surface air temperature is below the threshold (Eq. 7 and 8), these issues imply that both the precise amplitude of daily variation and accurate temporal patterns were particularly essential for cold damage forecasts.

Taking August 21–24 as an example of the overlapping period, the 1–4 days forecast for the cooling degree-days reduced 82.8% of RMSEs in the 11–14 days forecast (Fig. 9). The ratio generally increases along with the season; however, there was an exception found on August 11–14. This drop was caused by better performance in simulating surface air temperature throughout the period that initialized July 31 (Fig. 8a). The RMSEs of the cooling degree-days in the 11–14 days forecast was less increased from August 1–3; therefore, an improvement was suppressed on August 11–14. For the most cases, the increased effectiveness by forecast updating is attributed to both the surface air temperature and the DVI. Former analysis periods, such as the July 10 and July 20 initial date forecasts, are sufficient to count cooling degree-days, but the DVI did not reach 1.5 at this time; therefore, cold damage was uncounted (Fig. 6). In the July 31 initial forecasts, the condition of the DVI was satisfied, but the daily temperature was generally higher than the threshold, 22°C, in the analysis area. The forecasts initialized on August 10 and 20 satisfied both the conditions of temperature and DVI. Therefore, more cooling degree-days and a large improvement ratio were estimated compared to the forecasts starting from the other initial dates. This shows the importance of forecast updating increases in the cooling degree-days estimation with the season.

The spread to RMSE ratios indicate how well the spreads represent the uncertainties of forecasts (Takano, 2002). The ensemble spreads grew as the RMSEs increased throughout the forecast period. However, the spread to RMSEs ratios were smaller than 1, suggesting that the growth of ensemble spreads was insufficient for the ensemble system to represent uncertainties. The ratios for the cooling degree-days were smaller than those for the surface air temperature throughout the forecast period (Figs. 4, 5, and 8). Only the temperatures less than 22°C were considered to obtain the cooling degree-days. The smaller ratios for the cooling degree-days implied that the system had difficulty in representing the uncertainties of the forecasts under these conditions. The underdispersion of ensemble members in this system is an issue for future studies.

### 4.2 Synthesizing ensemble meteorological forest and crop growth simulations

The growth-based information would contribute to mitigating yield loss caused by abrupt temperature changes. When a daily temperature below 22°C is forecasted for one area, the occurrence of low-temperature sterility varies grid by grid because of different local growth stages. Our experiments show that the forecasts based on numerical weather predictions have seven days of predictability for surface air temperatures and at least two weeks for cooling degree-days. When we compare the performances to previous forecast systems, derived biases are similar or at a better level at the operational weather forecast system by the JMA for surface air temperature (JMA, 2019c). The RMSEs of daily maximum and minimum temperature for 2–7 days forecast were 2–3°C for the Tohoku region, and our RMSEs in the daily mean temperature was 1–2°C for the same forecast period (Fig. 7). This indicates a better performance in
simulating cooling degree-days than previous alert systems, however, it is difficult to evaluate its effectiveness because previous studies (NARO, 2019; JMA, 2019a) provided alerts in a qualitative manner and the performances of these systems are not currently opened. There are other previous studies that applied short or medium-range forecasts to agricultural sectors and showed the advantage of approximately one-week forecasts compared to seasonal forecasts; however, they focused on different factors to our simulation, such as farmers’ income (Roudier et al., 2016), economic value (Asseng et al., 2016), and evapotranspiration (Perera et al., 2014). Therefore, the RMSEs calculated in this study would be a reference value for relatively short-term forecasts in further studies.

4.3 Limitations of proposed ensemble forecast system for cold stress

The comparison of SIN and CLM forecasts suggests the feasibility of applying surface air temperature data produced by the numerical weather prediction system to forecasting the cooling degree-days. However, the surface air temperature forecasts include large errors, both biases and random errors, particularly in the northern part of the study area. The temperature fields over the area are largely influenced by the cool northeasterly wind and low-level clouds, implying that models have some difficulties in representing these conditions. To reduce the errors and improve cooling degree-days forecasts, it is essential to improve the numerical weather model, including the schemes for physical processes.

The spread to RMSE ratio indicated insufficient variances in the ensemble members in this system. Model biases can prevent the ensemble system from forecasting particular weather conditions. Improving the model should contribute to a better representation of uncertainties in the forecasts. Insufficient ensemble size also leads to ensemble members missing the possible states. The ensemble size of the system in this study was much smaller than that of the current operational ensemble prediction systems. Enhancing the ensemble size should help ensemble perturbations to capture the growing modes. Moreover, the system used in this study was based on a single set of atmospheric and rice growth models, so that the model uncertainties are not considered. Representing model uncertainties with multi-model ensemble method is also a possible approach to improving the relation between the ensemble spreads and the errors.

To apply the ensemble forecasting system to different regions, the phenology parameters should be turned to the cultivars cropped there, because this study focused on the specified cultivars cropped in northeastern Japan. As serious meteorological phenomena for rice growth vary by region, the predictability is not necessarily same as presented in this study. For applications of the forecasting system, users might consider regional differences in its performance and effectiveness.

5. Conclusion

We applied ensemble numerical weather prediction to rice growth simulations, focusing on the cold damage estimation in currently cold areas. The ensemble mean forecast prolonged the predictable period for two days, while the single deterministic forecast was superior to the climatological forecast for a FT of five days for surface air temperature. For cooling degree-days, while the climatological forecast simulated the rice growth stages reasonably, the forecasts based on numerical weather predictions performed better in all periods because of the better representation of daily variations in surface air temperature. Therefore, numerical weather prediction is becoming more important in the estimation of cooling degree-days.

Our simulation uses a very limited ensemble size of nine; therefore, further studies on the appropriate number of ensemble members would contribute to cold damage estimations. Increasing the ensemble size would propose more possible scenarios of cold damage risk, and could reduce the number of cases in which all ensemble members fail to predict cold damage. Moreover, more information on the reliability of the forecast would be provided, which helps farmers to avoid missing taking action. Both the intensity and reliability provided by the ensemble forecasting is valuable to take precautions against cold damage.

The application of numerical weather prediction data to estimate the yield of rice affected by temperature stress and external factors such as rice blast needs to be assessed in further studies. Taking appropriate action for daily variations is effective for both current daily precautions and future climate adaptations because of their similar order of variation.

Acknowledgements

This study was supported by JSPS KAKENHI (Grant Number: 16K18775, 18H04146 and 20K12191) and the Social Implementation Program on Climate Change Adaptation Technology (SI-CAT) Grant Number JPMXD0715667163 and the Integrated Research Program for Advancing Climate Models (TOUGOU) Grant Number JPMXD0717935561 from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. We thank Dr. Yasushi Ishigooka for providing the Mesh-AMeDAS observation dataset. Our thanks are extended to the editor and two anonymous reviewers for their fruitful comments.

References

Asseng S, McIntosh PC, Thomas G, Ebert EE, Kimashia N, 2016: Is a 10-day rainfall forecast of value in dry-land wheat cropping? Agricultural and Forest Meteorology 216, 170–176. Arague A, Vose RS, 2011: The definition of the standard WMO climate normal: The key to deriving alternative climate normals. Bulletin of the American Meteorological Society 92, 699–704. Easterling WE, Aggarwal PK, Batima P, Brander KM, Erda L, Howden SM, Kirilenko A, Morton J, Soussana J, Schmidhuber J, Tubiello FN, 2007: Food, fiber and forest products. In: Parry ML, Canziani OF, Palutikof JP., van der Linden PJ, Hanson CE, (eds) Climate change 2007: impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, p273–313. Fortin V, Abaza M, Ancil F, Turcotte R, 2014: Why should ensemble spread match the RMSE of the ensemble mean? Journal of Hydrometeorology 15, 1708–1713. Fukui S, Iwasaki T, Sha W, 2014: An ensemble downscaling
prediction experiment for medium-range forecast of the daily mean surface temperature distribution over northeastern Japan during summer. Journal of the Meteorological Society of Japan 92(6), 505–517.

Fukui S, Ishigooka Y, Kuwagata T, Hasegawa T, 2015: A methodology for estimating phonological parameters of rice cultivars utilizing data from common variety trials. Journal of Agricultural Meteorology 71(2), 77–89.

Hasegawa T, Ishimaru T, Kondo M, Kuwagata T, Yoshimoto M, Fukuoka M, 2011: Spikelet sterility of rice observed in the record hot summer of 2007 and the factors associated with its variation. Journal of Agricultural Meteorology 67(4), 225–232.

Horie T, 1988: The effects of climatic variations on agriculture in Japan. 5. The effects on rice yields in Hokkaido. In: Parry ML, Carter TR, Konjin NT (eds) The impact of climatic variations on agriculture. Vol 1: Assessments in cool temperature and cold regions. Kluwer, Dordrecht, p809–826.

Horie T, Nakagawa H, Centeno HGS, Kropff MJ, 1995: The rice crop simulation model SIMRiW and its testing. In: Matthews RB, Kropff MJ, Bachelet D, van Laar HH (eds) Modeling the impact of climate change on rice production in Asia. International Rice Research Institute and CAB International, Wallingford, p51–66.

Japan Meteorological Agency (JMA) 2019a: Examples of application for climate risk in agricultural sector (in Japanese). https://www.data.jma.go.jp/gmd/risk/taio_suitou.html, accessed 23 October 2019.

Japan Meteorological Agency (JMA) 2019b: climates in Tohoku region (in Japanese, https://www.jma-net.go.jp/sendai/wadai/kikouhena/column_tohoku.html, accessed 23 October 2019).

Japan Meteorological Agency (JMA) 2019c: Evaluation of weather forecast (in Japanese, https://www.data.jma.go.jp/fcd/yoho/kensho/yohohyoka_top.html, accessed 23 October 2019).

Kanda E, Kanno H, Okubo S, Shimada T, Yoshida R, Kobayashi T, Iwasaki T, 2014: Estimation of cool summer damage in the Tohoku region based on the MRI AGCM. Journal of Agricultural Meteorology 70(4), 187–198.

Kanno H, Watanabe M, Kanda E, 2013: MIROC5 predictions of Yamase (cold northeasterly winds causing cool summers in northern Japan). Journal of Agricultural Meteorology 69(3), 117–125.

Kobayashi T, Kanno H, Kanda E, Minamino K, Ardisyah POD, 2010: Development of rice climate damage reduction system using weather forecasting data based on Google Maps. Plant Protection 64(12), 794–800.

Klein WH, Lewis F, 1970: Computer forecasts of maximum and minimum temperatures. Journal of Applied Meteorology 9, 350–359.

Kojima M, Asano S, Iwabuchi H, 2006: Time-variations of optical and microphysical properties of Yamase clouds estimated from shipboard experiments and satellite remote sensing in June 2003. SOLA 2, 45–48.

Matsueda M, Nakazawa T, 2015: Early warning products for severe weather events derived from operational medium-range ensemble forecasts. Meteorological Applications 22, 213–222.

Ministry of Land, Infrastructure, Transport and Tourism (MLITT) 2012: National Land Numerical Information. MLITT, Tokyo.

Nakagawa H, Horie T, Matsui T, 2003: Effects of climate change on rice production and adaptive technologies. In: Mew TW, Brar DS, Peng S, Dawe D, Hardy B, (eds) Rice science: innovations and impact for livelihood. IRRI, Chinese Academy of Engineering, and Chinese Academy of Agricultural Sciences, p 635–658.

National Agriculture and Food Research Organization (NARO) 2019: Early warning system for paddy rice cold damage. (in Japanese; http://www.reigai.affrc.go.jp, accessed 23 October 2019).

Okada M, Iizumi T, Hayashi Y, Yokozawa M, 2011: Projecting climate change impacts both on rice quality and yield in Japan. Journal of Agricultural Meteorology 67(4), 285–295.

Perera KC, Western AW, Nawarathna B, George B, 2014: Forecasting daily reference evapotranspiration for Australia using numerical weather prediction outputs. Agricultural and Forest Meteorology 194, 50–63.

Ray DK, Ramankutty N, Mueller ND, West PC, Foley JA, 2012: Recent patterns of crop yield growth and stagnation. Nature Communications 3, 1293.

Roudier P, Alhassane A, Baron C, Louvet S, Sultan B, 2016: Assessing the benefits of weather and seasonal forecasts to millet growers in Niger. Agricultural and Forest Meteorology 223, 168–180.

Saito K, Ishida J, Aranami K, Hara T, Segawa T, Narita M, Honda Y, 2007: Nonhydrostatic atmospheric models and operational development at JMA. Journal of the Meteorological Society of Japan 85B, 271–304.

Seino H, 1993: An estimation of distribution of meteorological elements using GIS and AMeDAS data. Journal of Agricultural Meteorology 48, 379–383.

Shimada T, Kawamura H, 2011: summertime gap winds of the Soya Strait induced by the developed Okhotsk high. Atmospheric Science Letters 12, 316–320.

Shimono H, 2011: Earlier rice phenology as a result of climate change can increase the risk of cold damage during reproductive growth in northern Japan. Agriculture, Ecosystems & Environment 144, 201–207.

Shimono H, Kanno H, 2012: Risk map for cool injury inducing spikelet sterility by low temperature at booting stage in rice with taking into account for growth stage and temperature before the booting stage in Tohoku region, Japan. Climate in Biosphere 12, 6–11 (in Japanese with English abstract).

Simmons AJ, Hollingsworth A, 2002: Some aspects of the improvement in skill of numerical weather prediction. Quarterly Journal of the Royal Meteorological Society 128, 647–677.

Takanawa K, 2002: The application technique of ensemble forecasts. Meteorological Research Note 201, 73–103 (in Japanese).

Wassmann R, Jagadish SVK, Heuer S, Ismail A, Redona E, Serraj R, Singh RK, Howell G, Pathak H, Sunmpleth K, 2009: Climate change affecting rice production: the physiological and agronomic basis for possible adaptation strategies. Advances in Agronomy 101, 59–122.

Weisheimer A, Palmer TN, Doblas-Reyes FJ, 2011: Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. Geophysical Research Letters 38, L16703.

Yoshida R, Iizumi T, Nishinori M, 2012: Inter-model differences in the relationship between downward shortwave radiation and air temperatures derived from dynamical and statistical downscaling models. Journal of Meteorological Society of Japan 90B, 75–82.

Yoshida R, Fukui S, Shimada T, Hasegawa T, Ishigooka Y, Takayabu I, Iwasaki T, 2015: Adaptation of rice to climate change through a cultivar-based simulation: a possible cultivar shift in eastern Japan. Climate Research 64, 275–290.