# Monthly Weather Forecasts in a Pest Forecasting Context: Downscaling, Recalibration, and Skill Improvement

MARTIN HIRSCHI, CHRISTOPH SPIRIG, AND ANDREAS P. WEIGEL

*Federal Office for Meteorology and Climatology MeteoSwiss, Zurich, Switzerland*

PIERLUIGI CALANCA

*Agroscope Reckenholz-Tänikon Research Station, Zurich, Switzerland*

JÖRG SAMIETZ

*Agroscope Changins-Wädenswil Research Station, Wädenswil, Switzerland*

MATHIAS W. ROTACH

*Institute for Meteorology and Geophysics, University of Innsbruck, Innsbruck, Austria*

(Manuscript received 23 March 2012, in final form 29 June 2012)

ABSTRACT

Monthly weather forecasts (MOFCs) were shown to have skill in extratropical continental regions for lead times up to 3 weeks, in particular for temperature and if weekly averaged. This skill could be exploited in practical applications for implementations exhibiting some degree of memory or inertia toward meteorological drivers, potentially even for longer lead times. Many agricultural applications fall into these categories because of the temperature-dependent development of biological organisms, allowing simulations that are based on temperature sums. Most such agricultural models require local weather information at daily or even hourly temporal resolution, however, preventing direct use of the spatially and temporally aggregated information of MOFCs, which may furthermore be subject to significant biases. By the example of forecasting the timing of life-phase occurrences of the codling moth (*Cydia pomonella*), which is a major insect pest in apple orchards worldwide, the authors investigate the application of downscaled weekly temperature anomalies of MOFCs for use in an impact model requiring hourly input. The downscaling and postprocessing included the use of a daily weather generator and a resampling procedure for creating hourly weather series and the application of a recalibration technique to correct for the original underconfidence of the forecast occurrences of codling moth life phases. Results show a clear skill improvement of up to 3 days in root-mean-square error over the full forecast range when incorporating MOFCs as compared with deterministic benchmark forecasts using climatological information for predicting the timing of codling moth life phases.

## 1. Introduction

Output of the monthly weather forecast (MOFC) system of the European Centre for Medium-Range Weather Forecasts (ECMWF) is normally used in weekly aggregations, since the MOFC system does not have much skill to predict daily variability beyond 10 days of lead time (Vitart 2003; Vitart et al. 2008). For weekly averages, MOFCs show skill in extratropical continental regions for lead times up to 3 weeks, in particular for temperature (Weigel et al. 2008). In addition, the limit of predictability could even be extended beyond 3 weeks for users who only require forecast information on longer averaging intervals or for users who feed MOFCs into application models with some degree of memory or inertia (Weigel et al. 2008; Calanca et al. 2011). The common weekly format of MOFCs is often not suitable as direct input in application models, however. There are two major issues: 1) low spatial and temporal resolution and 2) the presence of significant biases. One option to address these issues is the application of statistical or dynamical downscaling

*Corresponding author address:* Martin Hirschi, Federal Office for Meteorology and Climatology MeteoSwiss, Krähbühlstrasse 58, 8044 Zürich, Switzerland.
E-mail: m.hirschi@alumni.ethz.ch

DOI: 10.1175/JAMC-D-12-082.1

(e.g., Calanca et al. 2009). Whereas dynamical downscaling is demanding in terms of computational resources, statistical downscaling can be achieved with relatively few resources and is thus attractive from an end-user perspective.

As opposed to seasonal forecasts (e.g., Hansen and Indeje 2004; Feddersen and Andersen 2005; Hansen et al. 2006), there are not yet many studies documenting the application of MOFCs. This study presents the use of MOFCs for forecasting agricultural pests by using the example of the codling moth (*Cydia pomonella*). An important objective is to show how the temporal and spatial mismatch between the aggregated information of the disseminated weekly MOFCs anomalies and the requirements of a real-world application can be resolved using standard downscaling tools. The codling moth is a major insect pest in apple orchards worldwide (Dorn et al. 1999), and damage is caused by the larvae, which burrow into the fruit to feed on the flesh. To avoid unnecessary treatments and with regard to sustainable plant-protection strategies, accurate forecasting tools have been developed to predict the development of codling moths depending on actual weather conditions. In Switzerland, such forecasts are operationally issued with the Schad-Organismen-Prognose auf Apfel (pest forecasting on apple; SOPRA) system (www.sopra.info; Samietz et al. 2007), which predicts the occurrence of codling moth life phases (and those of other insects) using hourly weather observations up to the day the forecast is issued and hourly climatological information (referred to hereinafter as "climatology") for the forecast range.

The specific aim of our study is to investigate the improvement in the skill of forecasting the timing of upcoming codling moth life phases using MOFCs instead of the hourly climatology. The spread in the MOFCs is thereby fully incorporated into the application model and allows for an uncertainty estimation of the forecast codling moth life-phase occurrences. For this purpose, we developed a statistical downscaling approach for MOFCs that is based on the combination of a stochastic weather generator and a subsequent resampling to translate weekly MOFC anomalies into suitable weather series to be used as input for pest forecasting.

## 2. Methods

### a. *Pest forecasts using hourly temperature sums*

The development of the investigated codling moth is highly determined by temperature. With temperature-sum-above-threshold models, the timing of the occurrence [as day of year (DOY)] of different codling moth life phases can be simulated (see Hirschi et al. 2012), most successfully by using the temperature (in hourly temporal resolution) of the immediate environment (habitat) of the corresponding life phase (Samietz et al. 2007). A particular life phase $x$ is reached once the following condition is fulfilled:

$$\sum_h \max(0, T_h - T_0) \geq S_{T_h,x},$$

where $T_h$ are hourly temperatures starting 1 January, $T_0 = 10°\mathrm{C}$ is the developmental zero of the codling moth postdiapause development (see, e.g., Howell and Neven 2000), and $S_{T_h,x}$ are location-specific and life-phase-specific sum thresholds. All relevant life phases starting with the flight start of the overwintering generation in spring are simulated, roughly spanning the time period from April to September.

### b. *Monthly weather forecasts and statistical downscaling approach*

MOFCs are issued once per week (i.e., one initialization per week; newly this is done 2 times per week; www.ecmwf.int/products/forecasts/d/charts/mofc_multi/forecast/) by ECMWF using the Variable Resolution Ensemble Prediction System (VarEPS; Buizza et al. 2007), which provides a 32-day outlook of temperature, precipitation, and other variables. The system is a coupled ocean–atmosphere general circulation model and applies an ensemble technique to sample initial-condition uncertainty (51-member ensemble). The MOFCs go along with 18 yr of hindcast data (5-member ensembles, covering 1993–2010 at the time of the analysis) that allow for the correction of systematic errors, for the implementation of statistical postprocessing techniques, and for systematic skill assessments. To enhance the robustness of the predictions and because the system does not have much skill on the daily time scale beyond 10 days of lead time (Vitart 2003), MOFCs are typically disseminated as weekly means for the upcoming 4 weeks (i.e., corresponding to days 5–11, 12–18, 19–25, and 26–32).

Because the temperature-sum-above-threshold models applied here require hourly input rather than weekly averages, a downscaling of the MOFC temperature anomalies was developed that includes the following three steps (Fig. 1a):

1) The first step is application of the "Met&Roll flexible and improved" (M&Rfi) parametric stochastic weather generator (WG) (Dubrovsky et al. 2004) to generate a large pool of daily weather series (i.e., 5000 yr of statistically consistent daily mean temperature and temperature range, solar radiation, and precipitation) representing the climate at a specific location. M&Rfi relies on a Markov chain to model precipitation occurrence, a gamma distribution to model the

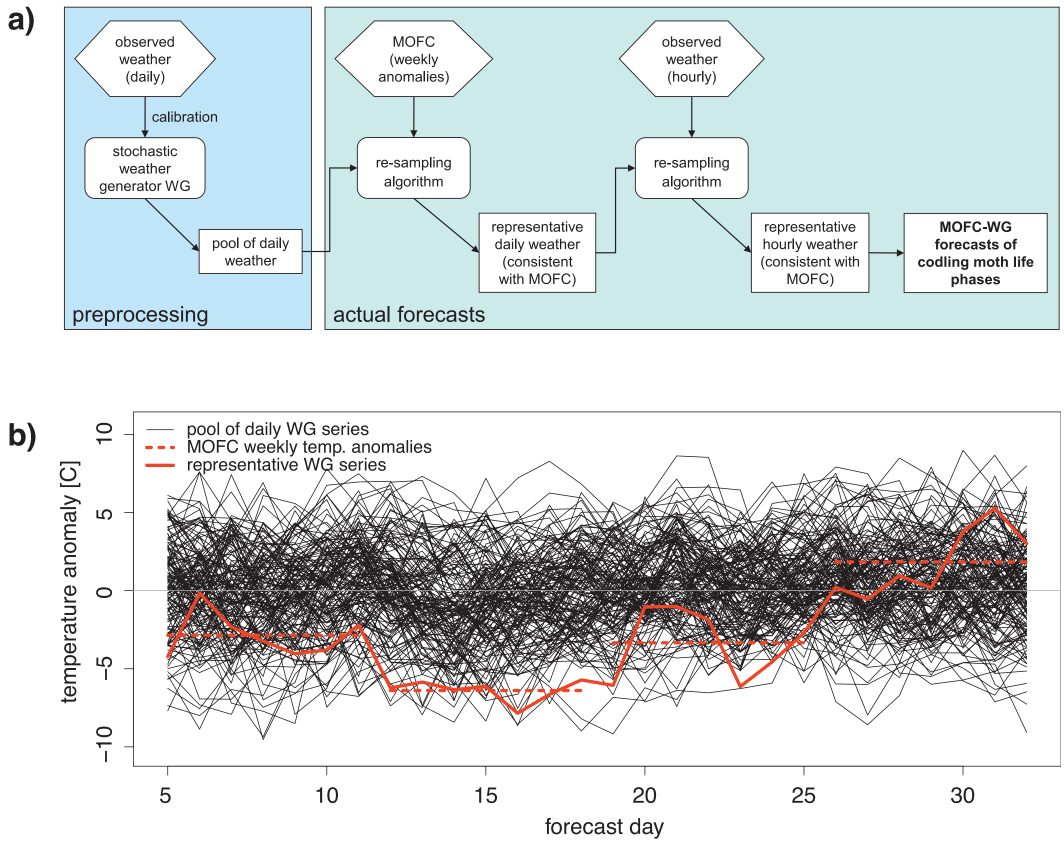


FIG. 1. Weather generator/resampling downscaling scheme: (a) overall setup and (b) resampling of a daily weather series from the pool of daily WG series to represent the weekly anomalies of one MOFC member.

precipitation amount, and an autoregressive model to simulate nonprecipitation variables (with conditioned statistics of the daily nonprecipitation variables on occurrence or nonoccurrence of precipitation). The stochastic WG was calibrated over the hindcast period of the MOFCs (i.e., 1993–2010) separately for each location (see below) by using daily meteorological observations.

2) The second step is resampling of representative daily weather series from this pool to optimally represent the evolution of the weekly anomalies of the individual MOFC members (Fig. 1b). This allows for a realistic representation of the location-specific day-to-day variability while integrating the week-to-week variation of the MOFC anomalies. The resampling is done on the basis of the weekly temperature anomalies using the scale-invariant Mahalanobis distance as the similarity measure (Mahalanobis 1936). Note that different weightings of the weeks and inclusion of precipitation as an additional variable in the resampling have been tested as well. These options showed no significant improvement in the skill of the pest forecasts as compared with the equal weighting of weekly temperature anomalies, but might be important for other applications.
3) The third step is extension of the daily weather series to hourly resolution on the basis of a resampling from hourly observations (similarity is again quantified by the Mahalanobis distance, considering daily mean precipitation and temperature, daily temperature range, and solar radiation; see also Hirschi et al. 2012).

The hourly temperature series consistent with the weekly MOFC anomalies were then used as input for the temperature-sum model of codling moth life phases to calculate pest forecasts once per week.

### c. *Experimental and verification setup*

The downscaling and forecasting procedure was evaluated on the basis of the simulated timing of dates of occurrence (expressed as DOY) of codling moth life phases using the temperature-sum forecasting model (section 2a) at the locations Wädenswil and Changins in the cooler northern part and Magadino in the warmer southern part of Switzerland. The dates were calculated using observed temperatures up to the last day before

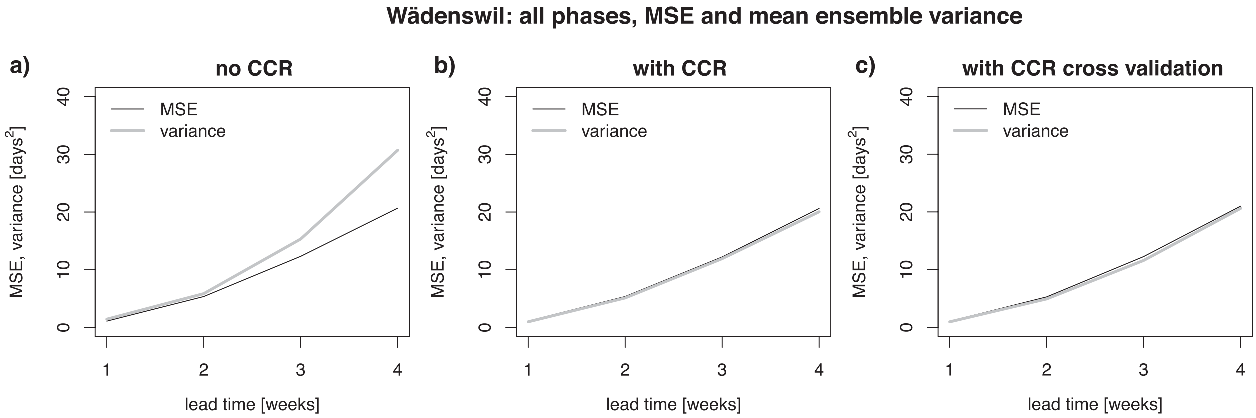


FIG. 2. MSE of the ensemble mean forecasts and time-mean intraensemble variance as a check for reliability of the MOFC-WG ensemble forecasts of codling moth life phases (a) before and (b) after application of the CCR, as well as (c) results of the recalibration when carried out in cross-validation mode (see text).

the first MOFC week and downscaled MOFCs thereafter. This "MOFC-WG" ensemble approach was compared with deterministic "benchmark" forecasts, using observed temperatures as above and climatological hourly means for the forecast range. Both MOFC-WG and benchmark were compared with the occurrences of life phases as derived with the temperature-sum model from observed temperatures (because of the lack of in situ observations of most codling moth life phases, this will be treated as truth and is denoted "reference").

From the MOFCs, 18 yr of hindcasts with five ensemble members for each year and the year 2011 with 51-member forecasts were downscaled and used for the MOFC-WG ensemble pest forecasts. The skill for predicting the timing of codling moth life-phase occurrences was evaluated at each of the four lead time weeks after forecast initialization. The lead times are labeled according to ECMWFs practice, for example, lead time week 1 (4) corresponds to days 5–11 (26–32) after forecast initialization.

For the evaluation of the pest forecast skill and for the comparison between the two forecast approaches (i.e., MOFC-WG ensemble forecasts vs deterministic benchmark), we applied scores that were calculated with respect to the reference occurrences of life phases (expressed as DOY) as derived from observed temperatures. Two scores were used:

1) The first score is the root-mean-square error (RMSE) of the MOFC-WG ensemble mean and the deterministic benchmark forecast.
2) The second score is the reliability of the MOFC-WG ensemble pest forecasts [REL; defined in Weigel et al.(2009), their Eq. (12)]. REL measures whether the ensemble members and the observed outcomes are consistently sampled from the same underlying probability distributions and thus are statistically indistinguishable from each other, implying that the ensemble spread is an appropriate estimate of the true forecast uncertainties (Weigel 2012). REL = 0 indicates reliable forecasts, REL > 0 indicates overconfidence, and REL < 0 indicates underconfidence.

## 3. Results and discussion

The hindcasts and 2011 forecasts of the timing of codling moth life-phase occurrences were evaluated for lead times of 1–4 weeks with respect to the reference occurrences of life phases as derived from observed temperatures.

### a. *Climate-conserving recalibration for the occurrence of codling moth life phases*

For all life phases considered and for weekly lead times 1–4, Fig. 2a shows the mean-square error (MSE) of the MOFC-WG ensemble mean forecasts and the time-mean intraensemble variance as defined by Weigel [2012, his Eq. (8.3) therein]. It is a necessary criterion for ensemble reliability that the MSE of the ensemble mean forecasts be equal to the time-mean intraensemble variance (Palmer et al. 2006; Weigel et al. 2009). The uncalibrated MOFC-WG hindcasts of codling moth life-phase occurrences show a clear tendency for underconfidence (i.e., intraensemble variance is too large; REL = −0.142, −0.046, −0.115, and −0.219 for lead times 1–4).

To overcome lack of reliability, ensemble forecasts can be recalibrated by using the climate-conserving recalibration (CCR; Weigel et al. 2009) technique, that is, by scaling the ensemble mean forecasts and the ensemble spreads with appropriate scaling factors derived from hindcasts and observed data. Here, CCR was

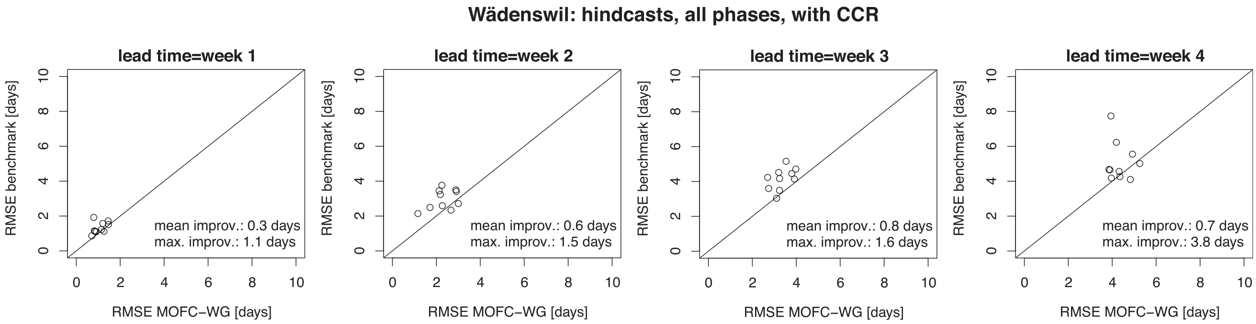


FIG. 3. Phase-specific RMSEs (i.e., of 10 life phases) of the forecast timing of codling moth life-phase occurrences in the hindcasts for lead times week 1–4: MOFC-WG ensemble mean vs climatological benchmark forecast. In each panel, the mean and the maximum improvement in RMSE (i.e., $RMSE_{benchmark} - RMSE_{MOFC-WG}$) over all phases is noted as well.

applied to the dates of occurrences of the pest life phases. The calibration factors were derived separately for each life phase to allow for seasonal dependence. This resulted in a successful correction of the underconfidence (Fig. 2b). For further investigation of the suitability of CCR for ensemble pest forecasts, a leave-one-out cross validation was carried out; that is, the recalibration of the hindcast of a specific year is based on the recalibration factors derived from the other years. This also led to a successful correction of the underconfidence of the original ensemble hindcasts (Fig. 2c).

### b. Skill improvement

Further verification showed a clear benefit of downscaled MOFC temperature anomalies for the full 4-week forecast range. Figure 3 presents the RMSEs of the hindcasts of all predicted life phases for both MOFC-WG and the climatological benchmark, itemized by the four lead time weeks. Likewise, the absolute deviations of the MOFCs ensemble mean and the benchmark in the 2011 forecasts are compared in Fig. 4. In the 2011 forecasts, not all phases were predicted at all lead-time weeks by the benchmark forecasts. For these cases, the absolute deviations of the corresponding MOFC-WG forecasts are plotted in the gray-shaded area of Fig. 4. The relatively large number of such events was a consequence of the observed large positive spring temperature anomaly in 2011 (see online at http://www.meteoschweiz.admin.ch/web/en/climate/climate_today/swiss_climate_maps.html), which caused an early occurrence of spring life phases in particular that could not be anticipated by the climatological forecasts. Note that, because of the fact that forecasts of a continuous variable (i.e., temperature) are translated into forecasts of a specific event (i.e., the timing of a particular life phase occurrence), the latter may be predicted by forecasts of more than one initialization or may be missed by the forecast at a particular lead time (as in the case of some benchmark forecasts in 2011). For the MOFC-WG ensemble forecast, this may result in varying ensemble sizes and has been taken into account in the verification (but is not further discussed here).

The hindcasts show a clear skill improvement of up to 3 days in RMSE over the full forecast range when incorporating MOFCs as compared with deterministic benchmark forecasts using climatology. Results for other skill measures (e.g., the probabilistic continuous rank probability score and the mean absolute error) as

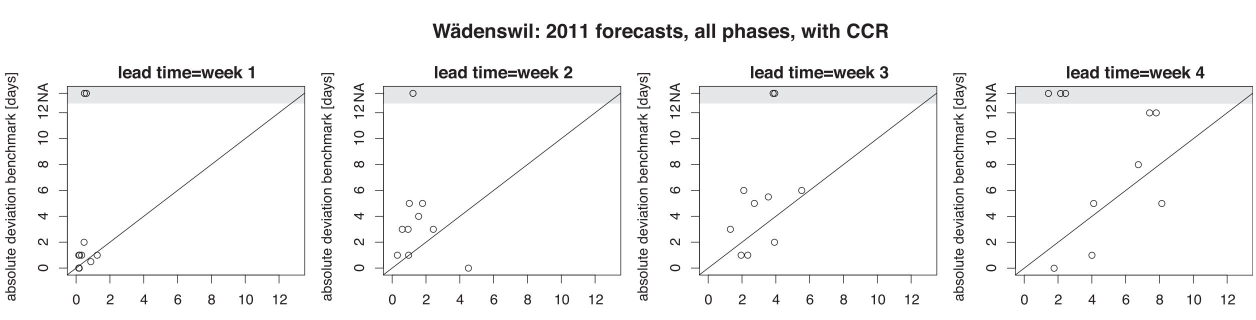


FIG. 4. As in Fig. 3, but for the 2011 forecasts and in terms of the absolute deviation. Life phases not predicted at a particular lead time by the benchmark forecasts are displayed in the gray-shaded area.

well as for the other locations exhibited a similar picture (not shown). The presence of forecast skill up to the full forecast range is promising. It demonstrates that skill can indeed be achieved significantly beyond 3 weeks of lead time (Weigel et al. 2008; Calanca et al. 2011) in this application, which involves some degree of memory because of the temperature-dependent development of the codling moth.

## 4. Conclusions

We demonstrated a way to apply weekly MOFC anomalies from ECWMF in a pest forecasting context requiring hourly input data. The applied statistical downscaling allows one to incorporate the information of MOFCs with relatively little effort for the end user. It implicitly removes local biases in the MOFCs and closes the gap between the common weekly MOFC format and the hourly temporal resolution required by the pest model. Similar downscaling approaches have been used in the past to translate seasonal or monthly forecasts for application models (e.g., Briggs and Wilks 1996; Wilks 2002; Calanca et al. 2011). The method applied here adds on to these studies by delivering hourly instead of daily data and furthermore by incorporating the temporal interdependence (i.e., the week-to-week variations) of the forecast anomalies. Overall, the inclusion of MOFCs improved the skill in forecasting the timing of codling moth life phases for all four weeks of the forecast range as compared with the climatological benchmark. Moreover, the incorporation of the spread of the MOFCs ensemble into the pest forecasting model additionally allows for an uncertainty estimation of the forecast codling moth life-phase occurrences.

*Acknowledgments.* We thank Martin Dubrovsky for providing the stochastic weather generator M&Rfi and for many helpful discussions. Funding was provided by the National Centre of Competence in Research on Climate (NCCR Climate) of the Swiss National Science Foundation.

REFERENCES

Briggs, W. M., and D. S. Wilks, 1996: Extension of the Climate Prediction Center long-lead temperature and precipitation outlooks to general weather statistics. *J. Climate,* **9,** 3496–3504.

Buizza, R., J.-R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart, 2007: The new ECMWF VarEPS (Variable Resolution Ensemble Prediction System). *Quart. J. Roy. Meteor. Soc.,* **133,** 681–695.

Calanca, P., L. Bogatai, T. Halenka, E. Cloppet, and J. Mika, 2009: Use of climate change scenarios in agrometeorological studies: Past experience and future needs. *Survey of Agrometeorological Practices and Applications in Europe Regarding Climate Change Impacts,* P. Nejedlik and S. Orlandini, Eds., COST ESSEM Publ., 237–266.

——, D. Bolius, A. P. Weigel, and M. A. Liniger, 2011: Application of long-range weather forecasts to agricultural decision problems in Europe. *J. Agric. Sci.,* **149,** 15–22, doi:10.1017/S0021859610000729.

Dorn, S., P. Schumacher, C. Abivardi, and R. Meyhöfer, 1999: Global and regional pest insects and their antagonists in orchards: Spatial dynamics. *Agric. Ecosyst. Environ.,* **73,** 111–118.

Dubrovsky, M., J. Buchtele, and Z. Zalud, 2004: High-frequency and low-frequency variability in stochastic daily weather generator and its effect on agricultural and hydrologic modelling. *Climatic Change,* **63,** 145–179.

Feddersen, H., and U. Andersen, 2005: A method for statistical downscaling of seasonal ensemble predictions. *Tellus,* **57A,** 398–408.

Hansen, J. W., and M. Indeje, 2004: Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya. *Agric. For. Meteor.,* **125** (1–2), 143–157.

——, A. Challinor, A. Ines, T. Wheeler, and V. Moron, 2006: Translating climate forecasts into agricultural terms: Advances and challenges. *Climate Res.,* **33,** 27–41.

Hirschi, M., and Coauthors, 2012: Downscaling climate change scenarios for apple pest and disease modeling in Switzerland. *Earth Syst. Dyn.,* **3,** 33–47.

Howell, J. F., and L. G. Neven, 2000: Physiological development time and zero development temperature of the codling moth (lepidoptera: Tortricidae). *Environ. Entomol.,* **29,** 766–772.

Mahalanobis, P., 1936: On the generalised distance in statistics. *Proc. Natl. Inst. Sci. India,* **2,** 49–55.

Palmer, T., R. Buizza, R. Hagedorn, A. Lawrence, M. Leutbecher, and L. Smith, 2006: Ensemble prediction: A pedagogical perspective. *ECMWF Newsletter,* No. 106, ECMWF, Reading, United Kingdom, 10–17.

Samietz, J., B. Graf, H. Höhn, L. Schaub, and H. Höpli, 2007: Phenology modelling of major insect pests in fruit orchards from biological basics to decision support: The forecasting tool SOPRA. *Bull. OEPP/EPPO,* **37,** 255–260.

Vitart, F., 2003: Monthly forecasting system. ECMWF Tech. Memo. 424, 70 pp.

——, and Coauthors, 2008: The new VarEPS-monthly forecasting system: A first step towards seamless prediction. *Quart. J. Roy. Meteor. Soc.,* **134,** 1789–1799.

Weigel, A. P., 2012: Ensemble forecasts. *Forecast Verification: A Practitioner's Guide in Atmospheric Science,* 2nd ed. John Wiley and Sons, 141–166.

——, D. Baggenstos, M. A. Liniger, F. Vitart, and C. Appenzeller, 2008: Probabilistic verification of monthly temperature forecasts. *Mon. Wea. Rev.,* **136,** 5162–5182.

——, M. A. Liniger, and C. Appenzeller, 2009: Seasonal ensemble forecasts: Are recalibrated single models better than multimodels? *Mon. Wea. Rev.,* **137,** 1460–1479.

Wilks, D. S., 2002: Realizations of daily weather in forecast seasonal climate. *J. Hydrometeor.,* **3,** 195–207.