CAMBRIDGE
UNIVERSITY PRESS

APPLICATION PAPER

Forecasting commodity returns by exploiting climate model forecasts of the El Niño Southern Oscillation

Vassili Kitsios[1,2,*] ⓘ, Lurion De Mello[3] and Richard Matear[4]

[1]Oceans and Atmosphere, CSIRO, Aspendale, Victoria, Australia
[2]Laboratory for Turbulence Research in Aerospace and Combustion, Department of Mechanical and Aerospace Engineering, Monash University, Clayton, Victoria, Australia
[3]Department of Applied Finance, Macquarie University, Sydney, New South Wales, Australia
[4]Oceans and Atmosphere, CSIRO, Hobart, Tasmania, Australia
*Corresponding author. E-mail: vassili.kitsios@csiro.au

Abstract

The physical and socioeconomic environments in which we live are intrinsically linked over a wide range of time and space scales. On monthly intervals, the price of many commodities produced predominantly in tropical regions covary with the dominant mode of climate variability in this region, namely the El Niño Southern Oscillation (ENSO). Here, for the spot prices returns of vegetable oils produced in Asia, we develop autoregressive (AR) models with exogenous ENSO indices, where for the first time these indices are generated by a purpose-built state-of-the-art general circulation model (GCM) climate forecasting system. The GCM is a numerical simulation which couples together the atmosphere, ocean, and sea ice, with the initial conditions tailored to maximize the climate forecast skill at multiyear timescales in the tropics. To serve as additional benchmarks, we also test commodity forecasts using: (a) no ENSO information as a lower bound; (b) perfect future ENSO knowledge as a reference upper bound; and (c) an econometric AR model of ENSO. All models adopting ENSO factors outperform those that do not, indicating the value here of incorporating climate knowledge into investment decision-making. Commodity forecasts adopting perfect ENSO factors have statistically significant skill out to 2 years. When adopting the GCM-ENSO factors, there is predictive power of the commodity beyond 1 year in the best case, which consistently outperforms commodity forecasts adopting an AR econometric model of ENSO.

Impact Statement

Combining skillful climate model forecasts of the El Niño/La Niña cycle, with econometric time series analysis methods, one can improve the ability to predict price changes in impacted commodities. This in turn assists producers, policymakers, and regulators to better anticipate and manage their future climate variability induced risk.

1. Introduction

The distinction between the economic impacts of weather and climate is somewhat simplistic, as it suggests that there are only two scales: (a) the local fast weather and (b) the global slow climate. In reality, both the human and geophysical systems are complex and vastly multiscale, comprising timescales

ranging from seconds to centuries. Geophysical phenomena have spatial scales ranging from waves that span the entire globe, down to millimeter-sized turbulence (Holton, 2004). Likewise, socioeconomic and market impacts are felt on a range of spatial scales from global factors of financial instability of the world's banking sector (Lamperti et al., 2019), down to the environmental hazard risk of individual farms (Fiedler et al., 2021). Despite the inherent nonlinear multiscale complexity of both the physical and economic systems, for certain spatiotemporal scales, there are potentially predictive relationships between the two.

Figure 1 represents the alignment between a subset physical phenomenon and economic agents. A selection of atmospheric (red), oceanic (blue), and coupled (gray) physical processes is illustrated with respect to their nominal spatiotemporal properties (Orlanski, 1975; Franzke et al., 2020; Ghil and Lucarini, 2020). To provide context, this figure is overlaid with approximate spatial scales of the places in which we live (e.g., continents and countries). Indicative investment time horizons are also included for various investors classes (e.g., hedge funds and central banks), which contribute to the multiscale nature of financial markets (Ghashghie et al., 1996; Mantegna and Stanley, 1996). Here, we address the continental scale multiyear economic impacts of climate variability, with specific application to forecasting vegetable oil markets. We exploit exogenous climate-related factors using standard econometric methods, to ascertain how the accuracy of climate predictions might influence the skill in the commodity forecasts.

The forecasting of the prices and yields of agricultural commodities has been studied for over a century (Moore, 1917; Allen, 1994). During this period, the use of time series analysis and econometric methods has dominated in forecasting research studies (Cuaresma et al., 2021). Orthogonal to the forecasting methods employed, the predictive power of various exogenous factors have also been assessed with application to the ability to forecast commodity prices. Numerous studies have focused on the potential predictive power of exchange rates and broader macroeconomic variables in forecasting commodity prices and indices (Chen et al., 2010; Groen and Pesenti, 2011; Gargano and Timmermann, 2014;
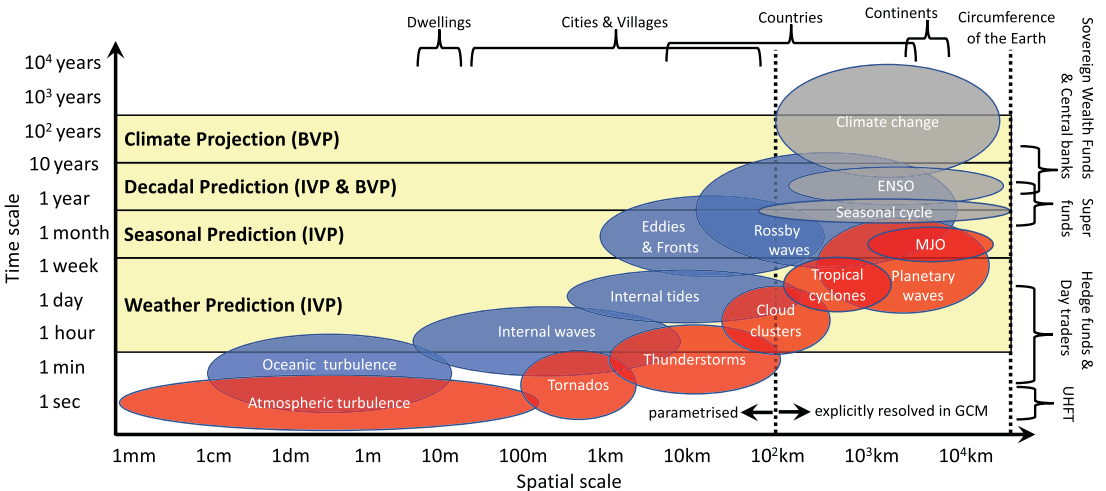

Figure 1. *Spatiotemporal scales of motion in the coupled Earth system, including atmospheric (red), oceanic (blue), and coupled (gray) physical phenomena. As lead time increases, the transition of forecasts being predominately an initial value problem (IVP) to a boundary value problem (BVP) is indicated by the yellow boxes. Using the areas (A) of dwellings, cities/villages, countries, and continents, we calculate the hydraulic diameters ($D = 2\sqrt{A/\pi}$) and overlay the figure with approximate ranges of D for each category along the spatial scale axis. The dotted black vertical line indicates spatial resolution of ocean grid in the climate reanalysis and forecast ensemble system.* Abbreviations: *ENSO, El Niño Southern Oscillation; GCM, general circulation model; MJO, Madden–Julian oscillation; UHFTs, ultra-high-frequency traders.*

Hollstein et al., 2021). As mentioned above, here we instead build models with exogenous climate-related factors using econometric methods. Note that the forecasting of an exogenous geophysical/environmental factor need not be undertaken in the same way as the forecasting of the commodity price itself. Various approaches to climate prediction are adopted herein, including both econometric methods, and also more computationally expensive physics-based dynamical simulations of the Earth system.

The development/specification of a numerically simulated geophysical forecasting system is intrinsically linked to the forecast time horizon of interest. Classes of such systems are illustrated by the yellow boxes in Figure 1. Climate projections are predominantly a boundary value problem, in that the determination of the system is primarily due to the specified boundary conditions of CO_2 and other forcings associated with prescribed scenarios of future economic development. The specifics of the initial conditions today are largely irrelevant for the properties of the Earth a century from now (with the possible exception of the slowest growing modes in the deep ocean). These types of simulations are most relevant for understanding the impacts of climate change under a given prescribed future scenario. Weather (lead time days) and seasonal (lead time months) prediction can be considered to be predominantly initial value problems, in which how well one knows the initial conditions is a key factor contributing to forecast skill. In decadal prediction (lead time years), both initial and boundary conditions are arguably important for many of the spatial scales (IPCC, 2013). When focused on making predictions for a given lead time, one only has the potential to predict the physical phenomena that persist over these associated timescales. Hence, in the context of making multiyear forecasts, the representation of physical phenomena with commensurate times scales is key. The El Niño Southern Oscillation (ENSO) is one such phenomenon as indicated by the gray ENSO oval in Figure 1. ENSO is characterized as an irregularly periodic variation in sea-surface temperatures and winds over the tropical regions in the Pacific Ocean (Bjerknes, 1969). It is the dominate mode of variability over the annual to multiyear timescale band.

ENSO has in fact played a significant role in the human response to climate, in particular with regard to agriculture, water, and health (Zebiak et al., 2015). One way to determine the economic risk associated with ENSO is to look at the relationship ENSO has with commodities and broader macroeconomic indicators. Macroeconomic growth can be affected by ENSO both directly impacting gross domestic product (GDP), and indirectly via its influence on the spillovers associated with trade and financial markets (Cashin et al., 2017). By exploiting any such relationships, one could potentially use ENSO forecasts to improve the prediction of market direction and risk of commodities and macroeconomic performance. For the period from the late 1950s to 1999, relationships were determined between ENSO events, nonoil primary commodity prices, measures of inflation, and growth GDP averaged across G7 countries (Brunner, 2002). ENSO variation was shown to have the largest impact on the prices of vegetable oils, grains, and some industrial commodities. Almost all of these commodities are significantly sourced from tropical regions, where the impact of ENSO is most direct. Coconut oil log returns were found to have the largest percentage contribution of ENSO to the forecast error variance at a 4-year lead time. El Niño was also shown to significantly affect the business cycles of South Africa, Australia, India, and arguably also Malaysia (Laosuthi and Selover, 2007). In general, countries that are most insulated from direct ENSO-related shocks are those that are geographically large, physically distant from the tropics, and with well-diversified economies (Laosuthi and Selover, 2007; Berry and Okulicz-Kozaryn, 2008). Hence, countries that are most sensitive to ENSO shocks are geographically small, located in the tropics, and with economic activities concentrated on few climate sensitive industries (Smith and Ubilava, 2017).

There have also been a wide range of studies on the relationships between ENSO and specific commodity yields and prices. Spatiotemporal relationships were demonstrated between sea-surface temperature anomalies in the equatorial Pacific and deviations in corn yields in the Midwest of the United States from long-term trends (Handler, 1990). Close linkages were found during the La Niña phase between soybean futures price movements and ENSO behavior for monthly timescales (Keppenne, 1995). U.S. wheat and corn futures were found not to have as significant relationships to ENSO, possibly due to institutional intervention. Negative (positive) correlations between sea-surface temperature anomalies and Brazilian corn production (prices) have also been found (Chimeli et al., 2008). The impact

on the phases of ENSO has also been quantified on the yields of soybean, wheat, rice, and maize (Iizumi et al., 2014). For wheat, rice, and maize, the neutral phase is the most productive, with increased magnitude in either phase of ENSO producing below average yield. For soybeans, El Niño produces above average yields, with yields in neutral and La Niña phases being below average. Focusing on the wheat market, a dynamic model was built comprising of the six main wheat-exporting countries, and the rest of the world as a single agent (Gutierrez, 2017). They showed that, on average, La Niña had a strong negative impact on wheat yield anomalies, exports, and stock-to-use ratios, and a positive impact on wheat export prices.

With application to the vegetable oil markets, nonlinear multivariate regression methods were used to identify asymmetric price responses to changes in ENSO (Ubilava and Holt, 2009). El Niño (La Niña) phases result in lower (higher) vegetable oil prices, with coconut and palm kernel oils more responsive to ENSO shocks than ground nut and cotton seed oils. The magnitude of the shock response on vegetable oil prices is larger if in the El Niño as opposed to the La Niña phase. A similar yet more broad analysis was performed, determining the predictive relationships between ENSO and 43 primary commodity spot prices from 1980 to 2016 (Ubilava, 2018). Commodities in the western region of the Pacific Ocean were found to respond most robustly to ENSO shocks, specifically protein meals and vegetable oils. Those with symmetric impacts of ENSO were shown to have improved out-of-sample forecast skill.

All of the aforementioned econometric studies have adopted some form of time series method for forecasting ENSO. In the present study, we determine if econometric forecasts of commodity spot prices can be further improved with the inclusion of exogenous ENSO factors provided by physics-based climate prediction simulations. The climate forecast dataset adopted within was generated using the climate reanalysis and forecast ensemble (CAFE) system (O'Kane et al., 2019), with its skill in forecasting ENSO previously assessed and compared to various similar systems in O'Kane et al. (2020). The commodities of interest are vegetable oils that are produced predominantly in the tropical regions of Asia, specifically coconut oil, palm oil, and soybean oil.

The manuscript is organized as follows. In Section 2, further background on ENSO is provided, along with details pertaining to the datasets, climate model forecasts, and autoregressive (AR) forecasting methods. In Section 3, the most parsimonious AR forecast model governing the evolution of ENSO is constructed on monthly intervals. The out-of-sample forecast skill of this AR model is then compared to that of ENSO climate model forecasts from the CAFE system. AR models are also built for the real log returns, and out-of-sample forecasts are undertaken for the vegetable oil commodities. The exogenous ENSO factors are provided by the CAFE system and also the AR model. To serve as additional benchmarks, we also test commodity forecasts using: (a) no ENSO information as a lower bound and (b) perfect future ENSO knowledge as a reference upper bound. We compare the skill of these cases, and in doing so demonstrate the potential for improved economic forecasts by exploiting climate predictions. Discussion linking these results back to the physical climate system is provided in Section 4, followed by concluding remarks and scope for future work in Section 5.

2. Background and Methods

A more detailed physical description of ENSO is first provided, followed by the associated observational and climate forecast datasets, and the time series analysis methods for forecasting both ENSO and the commodities.

2.1. Physical description of the El Niño Southern Oscillation

ENSO is a naturally occurring climate mode of variability with chaotic multiyear variations in sea-surface temperature and winds over the tropical Pacific Ocean (Bjerknes, 1969). Phenomenologically, it comprises of the mature phases La Niña and El Niño, and also the neutral phase. It takes between 3 and 7 years to transition from one mature phase to the other and back again, via the neutral state. A canonical La Niña is illustrated in Figure 2a, with cooler than average waters in the eastern Pacific (blue oval). It is warmer
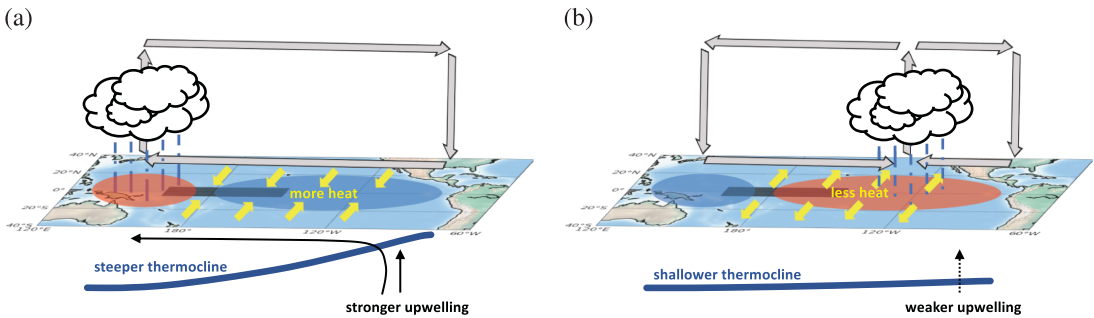
Figure 2. *Atmospheric and oceanic structure during phases of El Niño Southern Oscillation (ENSO): (a) La Niña and (b) El Niño. The black box indicates the region of the Niño4 index. The neural phase of ENSO is topologically similar to La Niña with less extreme anomalous temperature, winds, convection, and precipitation.*

than average in the west (red oval), which is accompanied by a local transfer of ocean mass to the atmosphere via enhanced evaporation. Ocean currents flow from the cooler water in the east, with an enhanced upwelling of the even cooler deeper water in the eastern Pacific. This means that the depth at which the ocean temperature matches that of the deep ocean shallows in the east due to upwelling, and deepens in the west due to anomalously warm surface temperatures. Along the equator, this depth is referred to as the equatorial thermocline. Typically, the thermocline slopes upward from west to east, and this slope is accentuated during a La Niña phase. This state has anomalously more heat (yellow arrows) transported to the region between the sea surface and the thermocline across the tropical Pacific. According to the recharge/discharge oscillator theory, this heat transfer plays a key role in the periodic transition between the phases of ENSO (Jin, 1997). In response to the ocean temperatures, the hot moist atmospheric air in the west convects upward, and this space is taken up by the cool dry air from the east. The hot air in the west travels toward the east where it cools, sinks to the surface, and completes an atmospheric loop referred to as the Walker circulation (gray arrows). As illustrated in Figure 2b, El Niño is characterized by an anomalously warm eastern Pacific (red oval), and cool western Pacific (blue oval). The slope of the thermocline decreases due to a deepening in the east and shallowing in the west, with less ocean heat content retained overall (yellow arrows). In the atmosphere evaporation, convection and rainfall are now enhanced in the central to eastern Pacific, and the Walker circulation breaks down. The neutral phase is topologically similar to La Niña, but with less extreme anomalous temperature, winds, convection, and precipitation. There is also evidence for a greater variety of transitional phases (Froyland et al., 2021). Due to climate teleconnections, ENSO can affect rainfall and temperature to more distant nonlocal regions of the Earth, with the strongest influences in the tropics. The phases of ENSO have seasonal, asymmetric, and heterogeneous impacts on rainfall and temperature (and potentially economic activity) over the entire Earth (Westra et al., 2015).

2.2. Climate observations

For the true historical observations of ENSO, we adopt the monthly averaged surface air temperature (SAT) fields from the Japanese reanalysis (JRA) dataset of Kobayashi et al. (2015). We calculate the phase and magnitude of ENSO using the typical Niño4 index, given by

$$E_{\text{obs}}(d) = \langle T_{0,\text{obs}}(x,y,d) - \overline{T}_{0,\text{obs}}(x,y,\tau) \rangle, \tag{1}$$

where $T_{0,\text{obs}}(x,y,d)$ is the monthly averaged SAT field at date d, of longitude x and latitude y. The angled brackets represent the spatial averaging within a tropical region in the eastern Pacific from 5°S to 5°N and 160°E to 150°W, represented by the black box in Figure 2. $\overline{T}_{0,\text{obs}}(x,y,\tau)$ is the climatological averaged SAT field, where τ represents the month of the year. A climatological average is the field (e.g., SAT)

averaged over an historical multiyear time period including only the desired month of the year (e.g., January and February). Here, this predefined historical period is set to be the in-sample period over which the model coefficients are learnt, which by definition is prior to the out-of-sample period for all forecasts presented in the current study. Positive values of $E_{\text{obs}}(t)$ indicate the El Niño phase, and negative values La Niña. The event strength is proportional to the magnitude of the index, with magnitudes less than half of the standard deviation (STD) of the signal typically classified as being in a neutral phase. In the out-of-sample forecasts, we use this dataset to calculate the ENSO-related model coefficients and initial conditions.

2.3. Climate model forecasting

As mentioned in Section 1, the general circulation model (GCM)-ENSO forecast dataset exploited here was generated using the CAFE system (O'Kane et al., 2019, 2020). The vast range of scales present in the global climate makes its numerical simulation exceedingly challenging. Even on the largest available supercomputers, it is not computationally tractable to have a grid domain large enough to capture the entire globe while having grid divisions small enough to resolve the finest scale turbulence. Since the representation of ENSO is key in this study, the adopted GCM must have sufficient spatial resolution to explicitly resolve its dynamics. This is illustrated in Figure 1, with the GCM ocean grid resolution indicated by the vertical dotted line being smaller than (or to the left of) the gray ENSO oval. Any physical phenomena to the left of the dotted line are unresolved with their influence on the large scales parameterized, and are hence not predictable with the current system. Specifically, the ocean grid has 1° resolution in longitude, higher latitudinal resolution in specific regions, and 50 vertical levels (Bi et al., 2013). The sea-ice model has the same horizontal resolution as the ocean grid, with five ice thickness categories. The atmospheric grid has a resolution of 2° in latitude, 2.5° in longitude with 24 vertical levels. The atmosphere is driven by realistic incoming solar radiation, aerosols, radiative gases, and land cover.

The climate forecasts are initiated from start dates at the beginning of each month between February 2002 and December 2015, inclusive. The initial conditions are tailored (or bred) to target the longer ENSO-related timescales, which comes at the expense of reduced skill at the shorter lead times (O'Kane et al., 2019). This point will become important in the discussion of the forecast skill metrics in the following sections, and is described in further detail in Appendix A. The associated Niño4 indices are calculated according to

$$E_{\text{gcm}}(d, \Delta t) = \langle T_{0,\text{gcm}}(x, y, d, \Delta t) - \overline{T}_{0,\text{obs}}(x, y, \tau) \rangle, \tag{2}$$

where $T_{0,\text{gcm}}(x, y, d, \Delta t)$ is the monthly averaged and ensemble-averaged SAT field from the GCM forecast dataset, Δt is the forecast horizon in months, and we again adopt the climatological average calculated from the JRA dataset $\overline{T}_{0,\text{obs}}(x, y, \tau)$ over a consistent climatological period. The GCM forecast initial conditions are indexed by d, where $d = 0$ refers to the first forecast initiated on February 1, 2002, $d = 1$ refers to March 1, 2002, and so on until the final forecast starting on December 1, 2015. Note that the GCM Niño4 index $E_{\text{gcm}}(d, \Delta t)$ is dependent upon both the initialization date d and the lead time Δt, since by virtue of the nonlinearity of the Earth system, a forecast 1 month ago will in general not coincide with the initial conditions today. The observations, however, are only dependent upon that date in question (i.e., $E_{\text{obs}}(d + \Delta t)$). In the econometric models of the commodities to follow, the monthly averaged climate data are only known at the beginning of the following month, to ensure that no future information is used.

These GCM Niño4 indices are additionally shifted and rescaled in a fair way using only data prior to the initialization data to remove persistent model biases, as detailed in Appendix B. Note that the debiasing requires several historical forecasts to be effective. For this reason, statistical measures of forecast skill are averaged over a common set of forecast start dates from February 2003 to December 2015.

2.4. Econometric time series forecasting of ENSO

We determine the most parsimonious AR representation of ENSO with additive seasonality, and compare its forecast skill to that of the GCM forecasts discussed in the previous section. We assume that the commodity price has no impact on ENSO, which is a typical assumption in the literature since ENSO is a naturally occurring climate mode that would have still been oscillating in the absence of humanity. As such, the ENSO model structure is given by

$$E(t) = a^{EE} + \sum_{l=1}^{11} a_l^{EE} D_l(t) + \sum_{k \in \mathbf{l}^{EE}} b_k^{EE} E(t-k), \tag{3}$$

where $D_l(t)$ is a binary variable equal to one during month l, and zero otherwise. The model offset a^{EE}, and associated month specific offsets a_l^{EE}, capture any potential additive seasonality. In the remaining summation, $\mathbf{l}^{EE}$ is a vector of lags with associated model gradients b_k^{EE}.

The procedure for determining the model structure is as follows. We assess all combinations of lags up to 12 months building models using all of the available data from January 1980 to December 2020. The first check is to ensure there is no significant serial autocorrelation in the residuals. Any models that contain residual lagged autocorrelations that exceed a 95% confidence level for any lag up to 12 months are excluded. Likewise, models that are shown to be heteroskedastic to a 99% confidence level on the basis of Engle (Engle, 1982) tests are also excluded. Of the remaining combinations of lags, we then determined the most parsimonious model as that minimizes the Bayesian information criterion (BIC; Schwarz, 1978)

$$\text{BIC} = 2N_p \log(N) + N \log\left(\frac{\text{RSS}}{N}\right), \tag{4}$$

where N is the number of samples, N_p is the number of parameters, and RSS is the residual sum of squares. We also assess the Akaike information criterion, its correction for small sample sizes, and the Hannan–Quinn information criterion, all yielding the same result for ENSO.

2.5. Econometric time series commodity forecasting

The most parsimonious AR representation of the commodity real log returns is determined as follows. The commodity prices ($P(t)$) used in this study were sourced from the World Bank database (www.worldbank.org/en/research/commodity-markets). The real commodity log returns $p(t) = (1 + i(t)) \times (1 + \widetilde{p}(t)) - 1$, where the nominal rate $\widetilde{p}(t) = \log(P(t)/P(t-1))$. The monthly inflation rate, $i(t) = \log(C(t)/C(t-1))$, is calculated on the basis of the G7 averaged consumer price index (denoted by $C(t)$), which was obtained from the OECD data portal (http://data.oecd.org).

Allowing for additive seasonality, AR processes, and lagged exogenous ENSO factors, the model representation of $P(t)$ is given by

$$p(t) = a^{pp} + \sum_{l=1}^{11} a_l^{pp} D_l(t) + \sum_{k \in \mathbf{l}^{pp}} b_k^{pp} p(t-k) + \sum_{k \in \mathbf{l}^{pE}} b_k^{pE} E(t-k), \tag{5}$$

where a^{pp} is the offset and a_l^{pp} the month-specific offsets. In the remaining summations above, $\mathbf{l}^{pp}$ is a vector of endogenous lags with associated model gradients b_k^{pp}, and $\mathbf{l}^{pE}$ is a vector of lags for ENSO influencing the evolution of the log returns in model gradients b_k^{pE}. The procedure for calculating the model coefficients is as discussed for the AR ENSO model; however, we now assess all combinations of lags up to 12 months in both the endogenous and exogenous variables.

3. Results

The AR- and GCM-ENSO forecasts are compared in Section 3.1, followed by an assessment of the commodity forecasts using various sources of ENSO information in Section 3.2.

3.1. ENSO forecasts

The most parsimonious model of ENSO in (3) has a lag structure of $\mathbf{l}^{ee} = (1, 4, 6)$. The minimum p-value for serial correlation is .142, the Engle test for heteroskedasticity returns a p-value of .854, and augmented Dickey–Fuller tests find no evidence for nonstationarity. All of which suggest that the residuals are stationary, white, and homoscedastic.

Both the AR-ENSO model and GCM forecasts are now assessed on the basis of out-of-sample forecast skill. For the AR model, the coefficients are calculated over a backward facing 19-year period ending at time $t - 1$, to ensure no future information is used in the forecast of start time t. As we move forward in time, the learning period also moves forward, and as such, the AR model parameters also vary. So, while the model structure is fixed, the model coefficients are a function of start date. Forecasts are made with lead times increasing in 1-month increments for the following 24 months, using the associated model parameters and requisite historical initial conditions. The debiased GCM forecasts are only available for a shorter period, with the common start dates from February 2003 to December 2015. The total number of samples per lead time used to calculate the forecast error metrics is 155.

Forecast skill metrics comparing the AR model and GCM forecasts are now assessed. In Figure 3a, the AR model has statistically significant positive correlations for the first 8 months, and insignificant correlations for the remainder of the window. The gray-shaded region in this plot, and all following correlation coefficient plots, indicate correlations that are statistically indistinguishable from zero to a 95% confidence level. The GCM forecasts have significant correlation for the entire 2-year forecast period. The AR model, however, outperforms the climate model for the first 4 months, with the climate model producing larger correlations for longer lead times. While the correlation coefficient indicates whether or not a set of forecasts is linearly proportional to the observations, it says nothing about the distance the forecasts are away from the observations. A better metric in this latter regard is the ratio of the root-mean-square error (RMSE) to the STD of the observations, which is illustrated in Figure 3b. A ratio less than 1 indicates that the RMSE is smaller than the natural variability of the system, and hence there is skill by this measure. This is the case for AR forecasts up to a lead time of 8 months, and for the GCM forecasts up to a lead time of 17 months. Again, the AR model outperforms the GCM in the earlier lead times with a normalized RMSE lower than the GCM up to a 5-month lead time.

Note that it is not guaranteed that a climate model will outperform an econometric forecast of the Niño index for all lead times. This is because the AR model is purpose-built to only forecast the Niño4 index, whereas the climate model must solve for the entire climate system at all points in time and space. As discussed in Section 2.3, the initial conditions in the GCM forecasts were designed to maximize their multiyear skill in the tropics, at the expense of skill at shorter lead times. These two sets of forecasts
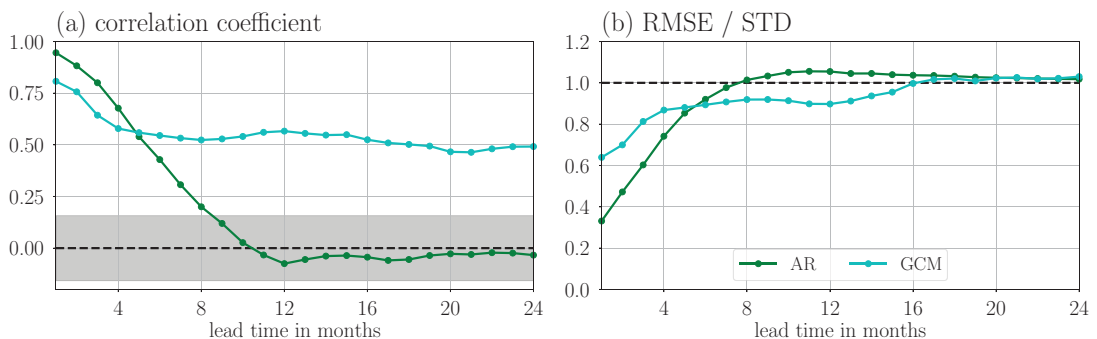

Figure 3. *El Niño Southern Oscillation forecast skill comparison of the out-of-sample error statistics for the AR model (green) and the general circulation model (cyan). (a) Correlation coefficient, with gray zone indicating statistically no different from zero to a 95% confidence level; (b) root-mean-square error divided by the standard deviation of observations, with legends applicable to both plots.*

(AR and GCM) will be sources of the exogenous ENSO factors in the commodity forecasts in the following section.

3.2. Commodity forecasts

The vegetable oil commodities of interest are: (a) coconut oil; (b) palm oil; and (c) soybean oil. These commodities have previously been found in the literature to have linear relationships with ENSO (Brunner, 2002; Ubilava, 2018). The endogenous lags of the log returns and exogenous lags of ENSO that minimize the BIC for each commodity are listed in Table 1, along with the minimum p-value for serial correlation (P_{serial}), and the Engle p-value for heteroskedasticity (P_{ARCH}). None of the presented models have statistically significant serial correlation to a 95% confidence level, nor heteroskedasticity to a 99% level, nor any evidence of nonstationarity on the basis of augmented Dickey–Fuller tests. We also consider the possibility that these commodities are additionally dependent upon global macroeconomic growth, by building more general AR models which also include the log returns of the G7 averaged GDP as an exogenous variable. None of these commodities contain GDP at any lag in their most parsimonious form. For completeness, the mathematical representation of the model incorporating GDP is detailed in Appendix C. To determine the sensitivity of the model structure to the data, we repeated the sweep of lags for all of the commodities using only historical data prior to the first forecast start date. In this instance, the optimal models have BIC values less than 1% smaller than the model structures deemed optimal when using all of the data.

The forecast correlation coefficients and normalized RMSE are illustrated for the commodities of coconut oil, palm oil, and soybean oil in Figure 4. The AR-ENSO case (solid green lines) represents the commodity forecasts with exogenous ENSO factors provided via the AR model of (3). As a test of the potential benefits of using the climate model forecasts, in the GCM-ENSO case (solid cyan lines), the CAFE system ENSO forecasts are adopted. As a potential upper bound on out-of-sample skill, in the perfect-ENSO case (solid blue lines), the ENSO forecasts are replaced with the actual future ENSO values. Note that the dashed blue line represents the perfect-ENSO case with model coefficients built using samples within which the forecasts are assessed, and as such are an upper bound on in-sample skill (or "predictive regressions" as referred to in the finance literature). Per start date, all of these cases have the same lag structure ($\boldsymbol{l}^{pp}$, $\boldsymbol{l}^{pE}$) and also the same values for the model coefficients. In the no-ENSO case (red line), however, the same endogenous lags ($\boldsymbol{l}^{pp}$) are adopted, but the model coefficients are recalculated such that this AR model best fits the data. The hollow black circles indicate lead times for which ENSO is Granger causal to a 95% confidence level on the basis of F-tests. That is, it determines if the RSS of the more complex commodity forecasts with additional ENSO factors is significantly lower than that of the simpler no-ENSO model, even after accounting for the additional number of parameters.

Coconut oil is perhaps the canonical example, illustrating how the adoption of higher fidelity ENSO forecasts can also improve the skill of the log returns. In Figure 4a, the correlation coefficient of the no-ENSO case has statistically significant positive correlations for only the first 4 months, while there is significant correlation for the entire 2-year period if using perfect future ENSO information with coefficients calculated from either future samples (dashed blue), or only historical data (solid blue). The fact that the solid and dashed blue lines have similar skill indicates that the relationships and hence model coefficients are consistent in time. These observations are echoed in the normalized RMSE plot in

Table 1. Autoregressive models of commodity real log returns with exogenous Niño4 factors.

Commodity	Endogenous commodity lags $\boldsymbol{l}^{pp}$	Exogenous ENSO lags $\boldsymbol{l}^{pE}$	P_{ARCH}	P_{serial}
Coconut oil	(1,3,12)	(8)	0.663	0.08
Palm oil	(1,2,4,9)	(6)	0.02	0.128
Soybean oil	(1,2,4)	(7)	0.891	0.504

Abbreviation: ENSO, El Niño Southern Oscillation.

(a) coconut oil

(b) coconut oil

(c) palm oil

(d) palm oil

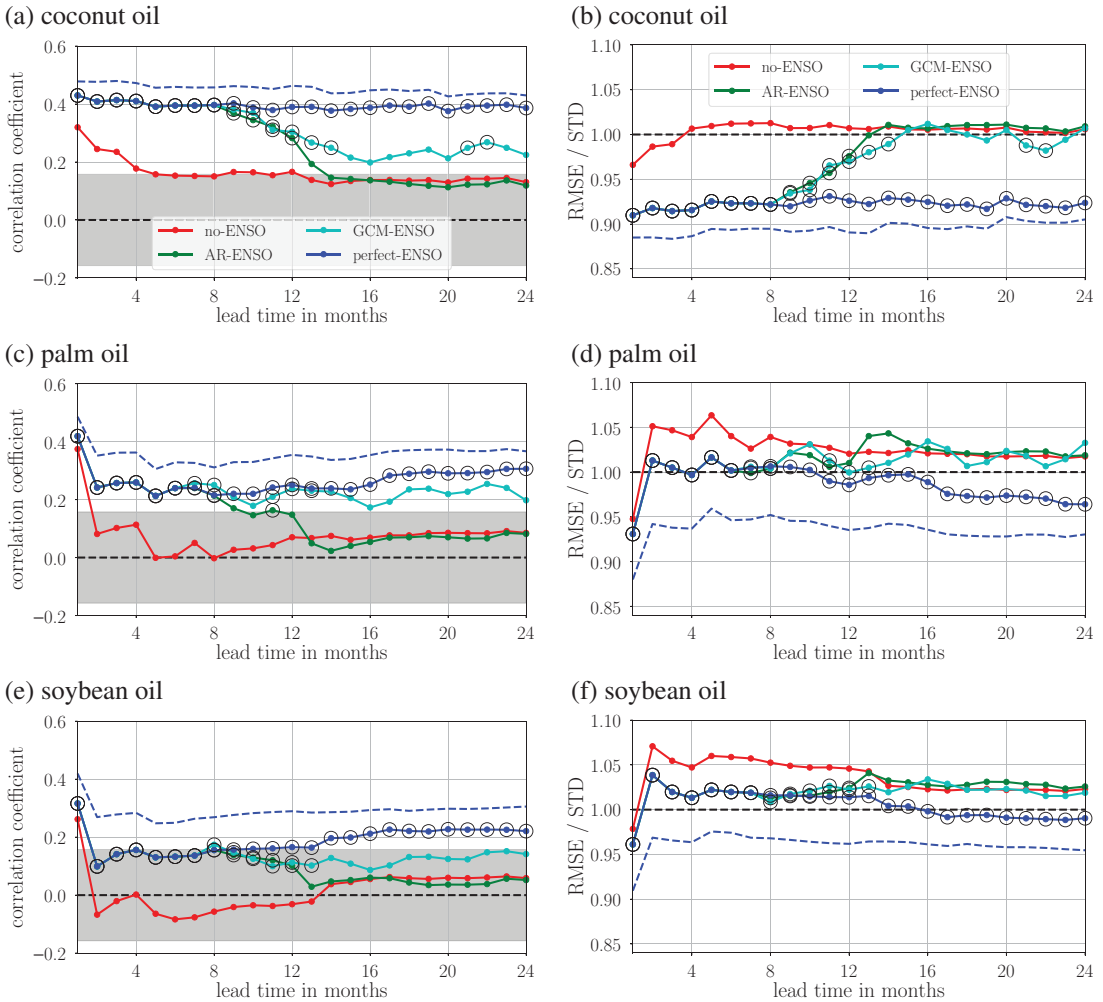### (e) soybean oil

(f) soybean oil

Figure 4. *Forecast skill comparison of the out-of-sample error statistics for the autoregressive (AR) models using no El Niño Southern Oscillation (ENSO; red), AR model of ENSO (green), general circulation model ENSO forecasts (cyan), and perfect future ENSO information (blue), on the basis of: (a) correlation coefficient for coconut oil, with gray zone indicating statistically no different from zero to a 95% confidence level; (b) root-mean-square error divided by the standard deviation of observations for coconut oil; (c) as in (a) but for palm oil; (d) as in (b) but for palm oil; (e) as in (a) but for soybean oil; and (f) as in (b) but for soybean oil. The hollow black circles indicate lead times at which VAR models with exogenous ENSO factors are found to be Granger causal in comparison to the no-ENSO case to a 95% confidence level. The dashed blue line represents the in-sample skill of the perfect-ENSO case with model coefficients calculated using all samples.*

Figure 4b, with the RMSE error less than natural variability only within the first 3 months for the AR model, and for the entire period in the perfect ENSO case. For the perfect-ENSO case, it is found that there is a significant reduction in RSS for all lead times (hollow symbols). This perfect-ENSO case, however, is not a realistic scenario since one cannot know ENSO perfectly into the future. The AR-ENSO case is a realistic set of forecasts, and still produces an improvement as compared to the no-ENSO case. Over the entire first year, it has both a statistically significant forecast correlation, an RMSE less than natural variability, and an RSS statistically significant less than the no-ENSO case. The GCM-ENSO case has

significant correlations for the entire 2-year forecast window, its RMSE is less than natural variability for the majority of the forecast horizon, and its RSS is significantly less than that of the no-ENSO case for the first 14 months of the forecast. The hollow circles at Months 21 and 22 should not be interpreted as a reemergence of causality or skill, but rather the RSS of the no-ENSO case increasing at a greater rate than that of the GCM-ENSO variant.

Over the first 8 months, the AR-ENSO, GCM-ENSO, and perfect-ENSO cases all have identical statistics. This is because the coconut oil commodity has an exogenous lag vector of $\mathbf{l}^{pE} = (8)$, as listed in Table 1. This means that up to this point, only historical ENSO information is required, and it is not until a lead time of 9 months that an actual ENSO forecast need be made. It is at this point that the forecast error measures begin to differ. Given that the AR model of ENSO outperforms the GCM for the early lead times (see Figure 3), one might then expect an advantage to adopting the AR-ENSO model in forecasting the commodities for the lead times immediately after 8 months. However, we find that whether AR-ENSO or GCM-ENSO model produces the better commodity forecasts depends on the lead time, skill measure selected, and commodity in question. This suggests that any initial difference in the ability of the AR model and the GCM to forecast ENSO is small in comparison to the error associated with the commodity forecasting model as a whole. The differences in the commodity price skill pertaining to the choice between the two ENSO forecasting methods are accentuated for longer lead times as the skill in the AR-ENSO variant drops off more rapidly with lead time.

There are qualitatively similar relationships for the other commodities. At any lead time in which the no-ENSO case has a significant correlation, it is consistently the worst performing case for all commodities on the basis of all error measures. For longer lead times in particular, following the perfect-ENSO case, the GCM-ENSO forecasts are the next most skillful, and then by AR-ENSO. The perfect-ENSO case has a statistically significant reduction in RSS (hollow circles) at all lead times for all commodities. It also has significant correlations into the second year of the forecasts, for all commodities. The GCM-ENSO case also produces statistically significant correlations into the second year of the forecasts for all commodities other than soybean oil.

4. Discussion

The reason why forecast skill of these particular commodities is improved with the inclusion of ENSO as an exogenous factor can be visualized via their lagged relationships with the physical climate system. As a reference, anomalous SATs averaged over La Niña and El Niño months are, respectively, presented in Figure 5a,b, with the dashed black lines the Niño4 region. These statistical representations of the phases of ENSO are consistent with the schematic depicted in Figure 2. Figure 5c illustrates the coconut oil log returns correlated with anomalous SAT anomaly lagged by 8 months, which was the optimal exogenous lag in the AR model. The correlation map clearly resembles the canonical El Niño in Figure 5b. This is also the case for palm oil and soybean oil each correlated with anomalous SATs shifted by the associated exogenous lags of 6 and 7 months, in Figure 5d,e. The countries in which these commodities are predominantly produced are indicated by the yellow stars in each of the maps, with: (a) coconut oil produced in Indonesia and the Philippines; (b) palm oil in Malaysia and Indonesia; and (c) soybean oil in China. The center for production of soybean oil, namely China, is the furthest physically removed from the tropics as compared to all of the aforementioned countries. Soybean oil also has the lowest magnitude in its correlation map over the Pacific ocean presumably due to its less direct influence from ENSO. Coconut oil and palm oil have more intense positive correlations in their respective maps over this region, hence demonstrating their strong relationships with ENSO. This is also consistent with these commodities having more improved skill due to the inclusion of the exogenous ENSO factors. Given that the correlation maps for all three commodities have spatial structures similar to the physical imprint of ENSO, it is unsurprising that physics-based climate models, which are better able to simulate and forecast ENSO, also contain more relevant information for forecasting quantities dependent upon ENSO.

It is interesting to note that the optimal ENSO-related lags are all of similar magnitude, that is, 8, 6, and 7 months for the coconut oil, palm oil, and soybean oil commodities, respectively. We speculate the
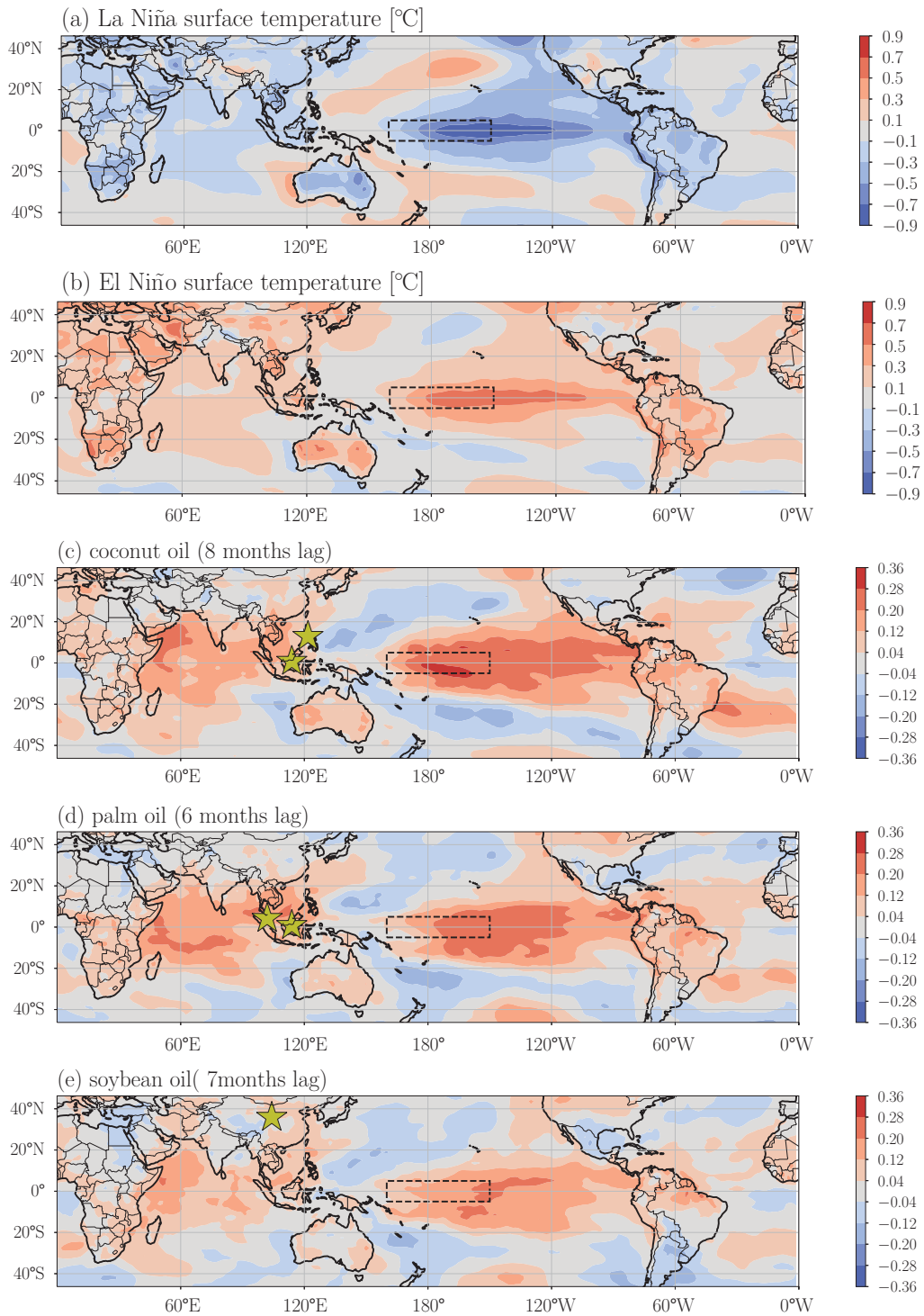
Figure 5. *Comparison of canonical El Niño Southern Oscillation and correlation patterns with commodity log returns. Surface air temperature (SAT) anomalies averaged over months in which the Niño4 index is at least half of a standard deviation: (a) below zero for La Niña; and (b) above zero for El Niño. The log returns of the commodities correlated with the SAT anomalies lagged by the number of months specified in the plot titles for: (c) coconut oil; (d) palm oil; and (e) soybean oil. Dashed Black line box is the Niño4 region. Yellow stars indicate centers of predominant production.*

following chain of influence. First, the distribution of sea-surface temperatures in the Pacific influences the atmospheric circulation and local climatic conditions where these crops are produced. Depending on how far removed the centers of production are from the equatorial Pacific, the local physical impact of ENSO could be delayed by a timescale in the vicinity of 1 month. The local climatic changes then impact the crop yields, which lead to a change in supply of that commodity to the market, flowing on to changes in price. Given this hypothesis, the supply chain duration involving crop planting, growth, harvesting, packaging, and logistics channels would set the ENSO-related lags. This suggests that developing forecast models incorporating crop yield and potentially supply chain factors explicitly may further assist in predicting these commodities.

5. Concluding Remarks

To recap, AR models were built for the real log returns of various commodity spot prices with the lagged ENSO index as an exogenous factor. All commodity forecasts adopting the ENSO factors outperformed optimal AR forecasts of the commodity that did not, indicating the value of incorporating climate information. For multiyear lead times, commodity forecasts adopting ENSO indices calculated from the GCM climate model were also found to be more skillful than those adopting an econometric model to predict the Niño4 time series. The greatest improvement in skill due to the inclusion of the ENSO factors was observed for the coconut oil market, followed by the palm oil market, with the soybean oil market having the least improvement. Skillful commodity return forecasts for coconut oil extend beyond 1 year when using dynamical climate forecasts of the ENSO. This is longer than our ability to forecast ENSO itself, which is due to the lagged relationship between ENSO and the commodity.

The additional utility of these physics-based ENSO forecasts is ultimately dependent upon the stakeholders' decision-making process. Concerning the coconut oil forecast skill, in comparison to the AR-ENSO variant, the GCM-ENSO forecasts offers 2 months of additional skill on the basis of the RMSE/STD metric, and 11 additional months on the basis of the correlation coefficient. If the direction and magnitude of the price movement is important, then the present GCM-ENSO forecasts offer an additional 2 months of potentially useful information. However, if the direction alone can contribute to making a better-informed decision, then the GCM-ENSO forecasts may be useful for an additional 11 months. The perfect-ENSO forecasts also provide an upper bound on the skill that one might be able to attain with an improved ENSO forecasting method. By this measure, there is scope for extended periods of utility if further advances in ENSO forecasts were achieved in the future, by whatever means (physics-based or otherwise). In general, the forecast structure presented within may be of practical use for stakeholders that are exposed to price movements in these commodities over multiyear timescales. Such stakeholders might include nations for which these commodities form a substantial part of their export revenue.

This study is particularly timely, given the recent increased prevalence of operational near-term climate forecasts provided by leading international centers as facilitated by the World Meteorological Organisation (Kushnir et al., 2019). Vast amounts of ensemble climate forecast data are being generated, adopting the latest advances in climate science. The current study is one demonstration of the utility in these physics-based multiyear climate forecasts to the financial sector in particular. One could also exploit ensemble forecasts of ENSO to produce an ensemble of associated commodity forecasts. This would enable not only the analyst to calculate the ensemble average as the best estimate of the forecast, but also the ensemble STD to quantify the commodity forecast uncertainty on the basis of the ENSO forecast uncertainty. Demonstrating the utility in forecasts of ENSO uncertainty is the subject of current research.

There are also some possible extensions that one might wish to implement in a production environment. First, for both the AR model of ENSO and that of the commodities, there are several lag combinations that have similar values of BIC, any of which would be defensible selections of the model structure. Here, we have chosen to use the optimal model in the forecasts, but one may also wish to run forecasts with an ensemble of model structures to quantify the model structure uncertainty. Second, for simplicity of communication, we calculated one time invariant model lag structure per commodity using

all of the available data. In a production environment, however, one may wish to instead reassess the lag structure each month using all of the historical data up to that point. Finally, using this methodology, one could also consider the influence of other climate teleconnections on relevant socioeconomic and market indicators.

References

Allen PG (1994) Economic forecasting in agriculture. *International Journal of Forecasting 10*, 81–135.

Berry B and Okulicz-Kozaryn A (2008) Are there ENSO signals in the macroeconomy? *Ecological Economics 64*(3), 625–633.

Bi D, **Marsland S**, **Uotila P**, **O'Farrell S**, **Fiedler R**, **Sullivan A**, **Griffies S**, **Zhou X and Hirst A** (2013) ACCESS-OM: The ocean and sea-ice core of the ACCESS coupled model. *Australian Meteorological and Oceanographic Journal 63*(1), 213–232.

Bjerknes J (1969) Atmospheric teleconnections from the equatorial pacific. *Monthly Weather Review 97*(3), 163–172.

Brunner A (2002) El Niño and world primary commodity prices: Warm water or hot air? *Review of Economics and Statistics 84*(1), 176–183.

Cashin P, **Mohaddes K and Raissi M** (2017) Fair weather or foul? The macroeconomic effects of El Niño. *Journal of International Economics 106*, 37–54.

Chen Y, **Rogoff K and Rossi B** (2010) Can exchange rates forecast commodity prices? *Quarterly Journal of Economics 125*, 1145–1194.

Chimeli A, **de Souza Filho F**, **Holanda M and Petterini F** (2008) Forecasting the impacts of climate variability: Lessons from the rainfed corn market in Cearà, Brazil. *Environment and Development Economics 13*(2), 201–227.

Cuaresma JC, **Hlouskova J and Obersteiner M** (2021) Agricultural commodity price dynamics and their determinants: A comprehensive econometric approach. *Journal of Forecasting 40*(7), 1245–1273.

Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica 50*(4), 987–1007.

Evensen G (2003) The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics 53*, 343–367.

Fiedler T, **Pitman A**, **Mackenzie K**, **Wood N**, **Jakob C and Perkins-Kirkpatrick S** (2021) Business risk and the emergence of climate analytics. *Nature Climate Change 11*, 87–94.

Franzke CLE, **Barbosa S**, **Blender R**, **Fredriksen H**, **Laepple T**, **Lambert F**, **Nilsen T**, **Rypdal K**, **Rypdal M**, **Scotto MG**, **Vannitsem S**, **Watkins NW**, **Yang L and Yuan N** (2020) The structure of climate variability across scales. *Reviews of Geophysics 58*(2), e2019RG000657.

Froyland G, **Giannakis D and Lintner B** (2021) Spectral analysis of climate dynamics with operator-theoretic approaches. *Nature Communications 12*, 6570.

Gargano A and Timmermann A (2014) Forecasting commodity price indexes using macroeconomic and financial predictors. *International Journal of Forecasting 30*, 825–843.

Ghashghie S, **Breymann W**, **Peinke J**, **Talkner P and Dodge Y** (1996) Turbulent cascades in foreign exchange markets. *Nature 381*, 767–770.

Ghil M and Lucarini V (2020) The physics of climate variability and climate change. *Reviews of Modern Physics 92*, 035002. https://doi.org/10.1103/RevModPhys.92.035002

Groen JJJ and Pesenti PA (2011) Commodity prices, commodity currencies, and global economic developments. In *Commodity Prices and Markets*, Vol. *20*. Chicago: University of Chicago Press, pp. 15–42.

Gutierrez L (2017) Impacts of El Niño Southern Oscillation on the wheat market: A global dynamic analysis. *PLoS ONE 12*(6), e0179086.

Handler P (1990) USA corn yields, the El Ninō and agricultural drought: 1867–1988. *International Journal of Climatology 10*(8), 819–828.

Hollstein F, **Prokopczuk M**, **Tharann B and Simen CW** (2021) Predictability in commodity markets: Evidence from more than a century. *Journal of Commodity Markets 24*, 100171.

Holton J (2004) *Dynamic Meteorology*. Burlington, MA: Elsevier.

Iizumi T, **Luo J-J**, **Challinor A**, **Sakurai G**, **Yokozawa M**, **Sakuma H**, **Brown M and Yamagata T** (2014) Impacts of El Niño Southern Oscillation on the global yields of major crops. *Nature Communications 5*, 3712.

IPCC (2013) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY: Cambridge University Press.

Jin F-F (1997) An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model. *Journal of the Atmospheric Sciences 54*, 811–829.

Keppenne C (1995) An ENSO signal in soybean futures prices. *Journal of Climate 8*(6), 1685–1689.

Kobayashi A, **Ota Y**, **Harada Y**, **Ebita A**, **Moriya M**, **Onoda H**, **Onogi K**, **Kamahori H**, **Kobayashi C**, **Endo H**, **Miyaoka K and Takahashi K** (2015) The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Series II 93*(1), 5–48.

Kushnir Y, **Scaife AA**, **Arritt R**, **Balsamo G**, **Boer G**, **Doblas-Reyes F**, **Hawkins E**, **Kimoto M**, **Kolli RK**, **Kumar A**, **Matei D**, **Matthes K**, **Müller WA**, **O'Kane T**, **Perlwitz J**, **Power S**, **Raphael M**, **Shimpo A**, **Smith D**, **Tuma M and Wu B** (2019) Towards operational predictions of the near-term climate. *Nature Climate Change 9*, 94–101.

Lamperti F, **Bosetti V**, **Roventini A and Tavoni M** (2019) The public costs of climate-induced financial instability. *Nature Climate Change 9*, 829–833.

Laosuthi T and Selover D (2007) Does El Niño affect business cycles? *Eastern Economic Journal 33*(1), 21–42.

Mantegna R and Stanley H (1996) Turbulence and financial markets. *Nature 383*, 587–588.

Moore HL (1917) *Forecasting the Yield and the Price of Cotton*. New York, NY: The Macmillian Company.

O'Kane T, **Sandery P**, **Monselesan D**, **Sakov P**, **Chamberlain M**, **Matear R**, **Collier M**, **Squire D and Stevens L** (2019) Coupled data assimilation and ensemble initialisation with application to multi-year ENSO prediction. *Journal of Climate 32*, 997–1024.

O'Kane T, **Squire D**, **Sandery P**, **Kitsios V**, **Matear R**, **Moore T**, **Risbey J and Watterson I** (2020) Enhanced ENSO prediction via augmentation of multi-model ensembles with initial thermocline perturbations. *Journal of Climate 33*, 2281–2293.

Orlanski I (1975) A rational subdivision of scales for atmospheric processes. *Bulletin American Meteorological Society 56*, 527–530.

Schwarz GE (1978) Estimating the dimension of a model. *Annals of Statistics 6*(2), 461–464.

Smith S and Ubilava D (2017) The El Niño Southern Oscillation and economic growth in the developing world. *Global Environmental Change 45*, 151–164.

Toth Z and Kalnay E (1997) Ensemble forecasting at NCEP: The breeding method. *Monthly Weather Review 125*, 3297–3318.

Ubilava D (2018) The role of El Niño Southern Oscillation in commodity price movement and predictability. *American Journal of Agricultural Economics 100*(1), 239–263.

Ubilava D and Holt M (2009) Nonlinearities in the world vegetable oil price system: El Niño effects. In *2009 Annual Meeting, July 26–28, 2009*. Milwaukee, WI: Agricultural and Applied Economics Association, pp. 1–27. http://doi.org/10.22004/ag.econ.49360

Westra S, **Renard B and Thyer M** (2015) The ENSO-precipitation teleconnection and its modulation by the interdecadal pacific oscillation. *Journal of Climate 28*(12), 4753–4773.

Zebiak S, **Orlove B**, **Muñoz A**, **Vaughan C**, **Hansen J**, **Troy T**, **Thomson M**, **Lustig A and Garvin S** (2015) Investigating El Niño–Southern Oscillation and society relationships. *Climate Change 6*, 17–34.

A. Climate Forecast Initialization Procedure

The initial conditions (or starting point) of the climate model forecasts are purpose-built in two phases. First, ensemble optimal interpolation (EnOI; Evensen, 2003) is used to correct the ocean of a single running climate model on the basis of available real-world surface and subsurface measurements of the ocean every 28 days (O'Kane et al., 2019). In a second phase, an ensemble of finite perturbation bred vectors (Toth and Kalnay, 1997) are then generated (or bred) about these corrected (or analyzed) fields. An ensemble of 10 additional fields is initialized by randomly perturbing the analyzed fields only within a volume surrounding the tropical ocean thermocline, where the dominant component of the temporal variability is in the 1–2-month timescale band (commensurate with the duration of a typical El Niño or La Niña event). Each of the perturbed members and the unperturbed control is marched forward in time for 28 days using the coupled GCM. The amplitude of the evolved anomalies with respect to the control is then rescaled to the finite amplitude of the original perturbations. This process is repeated using the next single EnOI analyzed state augmented with the ensemble of now evolved anomalies. Eventually, the random perturbations are forgotten, and physical structures appear in each of the bred vectors. The key point is that the initial conditions are tailored to the dynamics and timescales of interest. Refer to O'Kane et al. (2019) for further details on the breeding and rescaling procedure.

B. Climate Forecast Debiasing Procedure

The GCM Niño4 indices are additionally shifted and rescaled to remove persistent model biases. The debiased time series are given by

$$E_{\text{gcm}}(d, \Delta t) = \left[E_{\text{gcm}}(d, \Delta t) - \mu_{\text{gcm}}(d-1, \Delta t) \right] \frac{\sigma_{\text{obs}}(d-1+\Delta t)}{\sigma_{\text{gcm}}(d-1, \Delta t)} + \mu_{\text{obs}}(d-1+\Delta t), \text{where} \tag{6}$$

$$\mu_{\text{gcm}}(d, \Delta t) = \frac{1}{12} \sum_{i=d-11}^{d} E_{\text{gcm}}(i, \Delta t), \tag{7}$$

$$\sigma_{\text{gcm}}^2(d, \Delta t) = \frac{1}{12} \sum_{i=d-11}^{d} \left[E_{\text{gcm}}(i, \Delta t) - \mu_{\text{gcm}}(d, \Delta t) \right]^2, \tag{8}$$

$$\mu_{\text{obs}}(d) = \frac{1}{12} \sum_{i=d-11}^{d} E_{\text{obs}}(i), \text{and} \tag{9}$$

$$\sigma_{\text{obs}}^2(d) = \frac{1}{12} \sum_{i=d-11}^{d} [E_{\text{obs}}(i) - \mu_{\text{obs}}(d)]^2, \tag{10}$$

such that the above statistics are calculated over the most recent year. The fact that the averages and STDs in (6) are referenced at date $d-1$ ensures that only historical information is used to debias the forecasts of the start date d.

C. Significance Test for the Influence of Macroeconomic Growth on the Commodity Prices

Akin to what was done in Brunner (2002), we quantify global growth by the G7 averaged GDP (denoted below by $G(t)$), which was also obtained from the OECD data portal (http://data.oecd.org). Specifically, we test for the significance of any exogeneity from the real log returns of GDP, $g(t) = (1 + i(t)) \times (1 + \widetilde{g}(t)) - 1$, where the nominal rate $\widetilde{g}(t) = \log(G(t)/G(t-1))$. We undertake the same model building procedure articulated in Section 2.4, but now with respect to

$$p(t) = a^{pp} + \sum_{l=1}^{11} a_l^{pp} D_l(t) + \sum_{k \in V^{pp}} b_k^{pp} p(t-k) + \sum_{k \in V^{pg}} b_k^{pg} g(t-k) + \sum_{k \in V^{pE}} b_k^{pE} E(t-k), \tag{11}$$

where in the above summations $V^{p\eta}$ is a vector of lags for exogenous variable η influencing the evolution of the endogenous variable p, with associated model gradients $b_k^{p\eta}$. The global offset a^{pp}, and associated month specific offsets a_l^{pp}, capture any potential additive seasonality. This model is equivalent to that in (5), but with an additional summation to capture any lagged exogeneity from $g(t)$. Note that this analysis is undertaken on quarterly intervals, since the GDP data are only available per quarter. All combinations of lags are assessed up to three quarters. On the basis of F-tests, we calculate p-values for the significance of the reduction in the residual sum of squares as a result of including GDP as an exogenous variable for all lead times up to 2 years.