Atmospheric Precursors of Skillful SST Prediction in the Northeast Pacific

A. A. CLUETT,[a,b] M. G. JACOX,[a,c] D. J. AMAYA,[c] M. A. ALEXANDER,[c] AND J. D. SCOTT[c,d]

[a] *NOAA/Ecosystem Sciences Division/Southwest Fisheries Science Center, Monterey, California*
[b] *Fisheries Collaborative Program, Institute of Marine Sciences, University of California Santa Cruz, Santa Cruz, California*
[c] *NOAA/Physical Sciences Laboratory/Earth System Research Laboratory, Boulder, Colorado*
[d] *Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado*

(Manuscript received 21 February 2024, in final form 8 July 2024, accepted 10 July 2024)

ABSTRACT: Forecasts of sea surface temperature anomalies (SSTAs) provide essential information to stakeholders of marine resources in coastal ecosystems, such as the California Current Large Marine Ecosystem (CCLME), at management-relevant monthly-to-annual time scales. Diagnosing dynamical sources of predictability and the mechanisms differentiating skill among forecasts is required for verification and improvement in operational forecasting systems. Using retrospective forecasts (1982–2020) from a four-member subset of the North American Multi-Model Ensemble (NMME), we evaluate the conditional skill of SSTA forecasts in the CCLME at monthly resolution for lead times up to 10.5 months. Forecasts from ensemble members with relatively small SSTA errors at shorter lead times retain higher skill at longer lead times, with the most substantial and long-lasting increases for forecasts initialized in the fall and early spring. The "best" low-error SSTA forecasts are characterized by increased skill in the prediction of North Pacific atmospheric circulation [sea level pressure (SLP) and 200-hPa geopotential height] the month prior to the evaluation of SSTA errors in the CCLME and exhibit more realistic progressions of anomalous SLP. The Pacific meridional mode (PMM) emerges as a diagnostic of skillful North Pacific atmosphere–ocean coupling, as forecasts that correctly simulate the PMM and its associated SLP variability increase the SSTA prediction skill in the CCLME in the fall through spring. Predictable coupled ocean–atmosphere modes provide a target for enhancing predictability with early detection of the onset of a deterministic progression emerging from stochastic atmospheric variability.

SIGNIFICANCE STATEMENT: Global forecast systems provide near-term climate predictions that inform the management of marine resources, such as those of the California Current Large Marine Ecosystem. In this study, we probe the processes which lead forecasts to succeed or fail at predicting sea surface temperatures in the California Current at seasonal time scales among retrospective forecasts from the North American Multimodel Ensemble. We demonstrate that forecasts which best simulate sea surface temperatures at the earliest lead times sustain advantages in forecast skill and find that correctly simulating extratropical atmospheric circulation increases the predictive skill of sea surface temperatures in the northeast Pacific in the following lead times. Our results offer North Pacific atmospheric circulation as a target for forecast model improvement that would additionally enhance ocean forecasts.

KEYWORDS: North Pacific Ocean; Sea surface temperature; Forecast verification/skill; Seasonal forecasting

1. Introduction

The California Current System (CCS) is an ecologically and socioeconomically important large marine ecosystem (LME) in the coastal northeast Pacific, which supports diverse harvested and protected species (Checkley and Barth 2009; Sherman and Alexander 1986). Integration of climate prediction into dynamic management of such living marine resources enables proactive decision-making by managers and other stakeholders (Eveson et al. 2015; Tommasi et al. 2017), provided that forecasts accurately convey necessary physical or biogeochemical variables [e.g., sea surface temperature (SST)] at time horizons relevant to both decision-making and predictability (Brady et al. 2020; Doblas-Reyes et al. 2013; Hobday et al. 2013, 2018; Jacox et al. 2020; Yeager et al. 2018). As modern climate variability departs from historical norms, diagnosing, verifying, and harnessing the dynamical sources of skill underlying SST prediction present a path toward more confident forecasting in the California Current LME (CCLME) for operational usage and an opportunity to identify targets for improved simulation of the North Pacific and global climate system.

Global forecast systems, which are more readily available than downscaled regional products, exhibit significant skill in predicting coastal SST anomalies (SSTAs) in the CCLME despite their relatively coarse horizontal resolution (Hervieux et al. 2019; Jacox et al. 2019; Stock et al. 2015). The North American Multi-Model Ensemble (NMME) is a suite of state-of-the-art global climate forecast models run in real time and retrospectively (Becker et al. 2022; Kirtman et al. 2014). By gathering models with different formulations of physics,

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-24-0121.s1.

Corresponding author: Allison Cluett, allison.cluett@noaa.gov

DOI: 10.1175/JCLI-D-24-0121.1

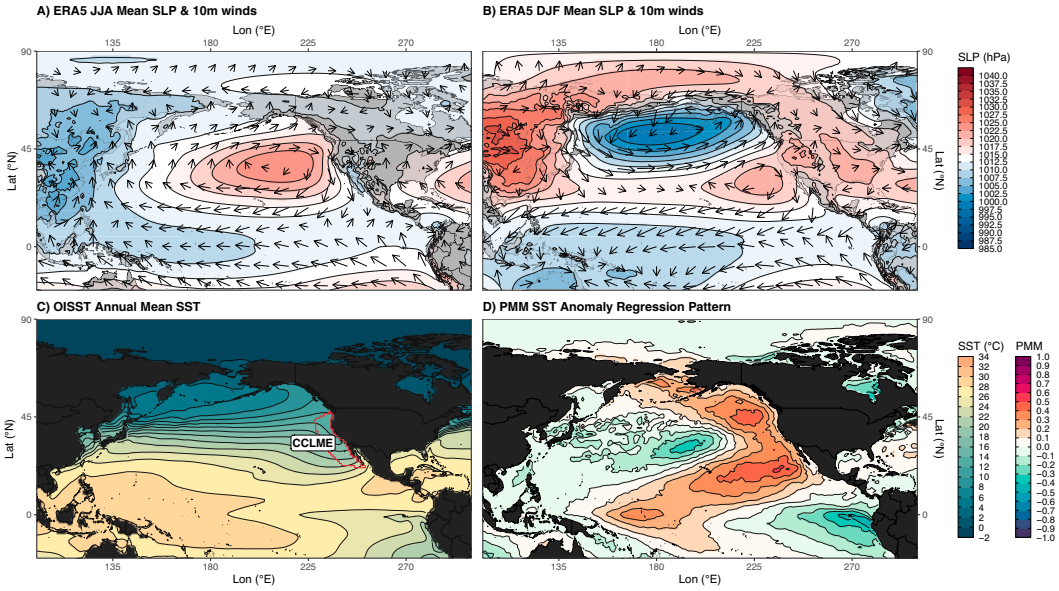

FIG. 1. Maps of study domain. Mean North Pacific atmospheric circulation, during (a) boreal summer (JJA) and (b) winter (DJF). Contour lines and color scale indicate pressure reduced to mean sea level, and arrows indicate 10-m winds calculated from *U* and *V* vectors from the ERA5 reanalysis monthly means for 1982–2020. (c) Climatological mean SSTs from 1982 to 2020 from the NOAA OISSTv2 product, with contour lines at 2°C and the CCLME region outlined in red. (d) Regression pattern of SST anomalies on the PMM, calculated from the OISSTv2 for 1982–2020.

ensembling strategies, initialization methods, and native resolutions, multimodel systems such as the NMME reduce the biases and errors associated with a single modeling approach (e.g., Becker et al. 2022). An additional ensemble of forecasts within each NMME model encompasses uncertainty in initial conditions. Although the multimodel ensemble mean (EM) forecast is used most frequently and often performs best at prediction of SSTA (Jacox et al. 2019; Kirtman et al. 2014; Stock et al. 2015), subsets of ensemble members outperform the ensemble mean in some contexts (Hervieux et al. 2019).

In the northeast Pacific, large-scale atmospheric forcing is a leading driver of SST variability from monthly to multidecadal time scales (Alexander et al. 2002; Chhak et al. 2009; Deser et al. 2010; Johnstone and Mantua 2014; Miller et al. 1994; Trenberth and Hurrell 1994). In the summer, the anticyclonic flow around a broad North Pacific High drives persistent northerlies along the west coast of North America, promoting upwelling of cold, nutrient-rich waters throughout the CCLME (Figs. 1a,c) (Checkley and Barth 2009). In the winter months, the Aleutian low (AL) develops and intensifies the cyclonic flow north of 45°N, strengthening midlatitude westerlies, weakening equatorward winds along the west coast, and reducing upwelling in the CCLME (Jacox et al. 2018) (Fig. 1b). Because regional-scale skill in global forecast systems derives from sensitivity to basin-scale variability (Stock et al. 2015), predictability of SSTA in the CCLME likely, in large part, results from linked variability in large-scale atmospheric circulation and these regional-scale wind-driven ocean processes (e.g., upwelling, horizontal advection, and vertical mixing) (Jacox et al. 2019; Stock et al. 2015). However, atmospheric forcing of ocean processes displays high-frequency stochastic variability, which drives dispersion among forecasts and limits the time scales of predictability for variables such as sea level pressure (SLP) and near-surface winds (Chikamoto et al. 2015; Davis 1976; Liu and Di Lorenzo 2018).

Dynamic ocean processes can integrate high-frequency atmospheric variability and, as if through a low-pass filter, yield SSTAs that feedback to the atmosphere (Di Lorenzo et al. 2013; Frankignoul and Hasselmann 1977). Moreover, coupled ocean–atmosphere interactions and related teleconnections can produce predictable spatiotemporal variability at extended time scales in the North Pacific (Capotondi et al. 2019; Vimont et al. 2003; Di Lorenzo et al. 2008). For example, substantial skill in the monthly SSTA prediction in the CCS arises in connection to El Niño–Southern Oscillation (ENSO), as ENSO-related predictability is propagated from the tropics by atmospheric teleconnections and oceanic coastally trapped waves (Hervieux et al. 2019; Jacox et al. 2019; Stock et al. 2015; Alexander et al. 2002). The Pacific meridional mode (PMM) is the second leading mode of ocean–atmosphere variability in the North Pacific (Chiang and Vimont 2004), distinct from—though likely interacting with—ENSO (Vimont et al. 2003; Amaya 2019; Larson and Kirtman 2014; Thomas and Vimont 2016). The PMM describes a gradient between anomalous warmth in the north tropical Pacific and the eastern equatorial Pacific cold tongue (Chiang and Vimont 2004) but displays SSTA extending into subtropics and the northeast Pacific throughout the CCLME and has been identified as an important precursor to large SSTA along the U.S. West Coast (Amaya 2019; Capotondi et al. 2019) (Fig. 1d). Anomalous SLP and winds related to the North Pacific Oscillation (NPO) lead the full expression

TABLE 1. Forecast models and verification datasets used.

Model/dataset	Ensemble	Organization	Reference
CanCM4i	10	Canadian Meteorological Center (CMC)	Merryfield et al. (2013)
GEM-NEMO	10	Recherche en Prévision Numérique (RPN)	Lin et al. (2019)
NCEP-CSFv2	24/28	National Centers for Environmental Prediction (NCEP)	Saha et al. (2014)
COLA-RSMAS-CCSM4	10	National Center for Atmospheric Research (NCAR)	Infanti and Kirtman (2016)
		Verification data	
OISSTv2	—	National Oceanic and Atmospheric Administration (NOAA)	Banzon et al. (2016)
ERA5	—	European Center for Medium-Range Weather Forecasting (ECMWF)	Hersbach et al. (2020)

of the PMM SSTA footprint by at least 1 month, serving as a stochastic trigger to a deterministic sequence of air–sea interactions that unfold most prominently in the winter and spring (Chiang and Vimont 2004).

In this study, we investigate the role of atmospheric forcing and PMM-related ocean–atmosphere interactions in driving SSTA skill in the CCLME at seasonal (monthly-to-annual) time scales. To understand the contributions of skill associated with specific processes, we calculate conditional skill, which describes the dependence of skill for one variable at a given lead time on the skill of that same variable (or another) at a different lead time (Dias and Kiladis 2019). Calculation of conditional skill enables the assessment of if and when forecast skill at short lead times is a reliable predictor of forecast skill at longer lead times, thereby revealing relationships in skill along the evolution of a forecast. Specifically, we address the following questions in a series of forecast-conditioning experiments: 1) Do forecasts with low errors in SSTA at short lead times sustain an advantage in predictive skill at longer lead times? 2) Are there atmospheric and/or oceanic precursors associated with skillful SSTA forecasts in the CCLME? and 3) How do atmospheric processes and coupled modes contribute to spatiotemporal patterns in predictability and prediction skill of SSTA in the CCLME and North Pacific?

2. Methods

a. *Forecast models and verification data*

We analyze a four-model subset (CanCM4i, GEM-NEMO, NCEP-CFSv2, and COLA-RSMAS-CCSM4; Table 1) of the NMME (Kirtman et al. 2014) Phase II, accessed through the IRI/LDEO database (https://iridl.ldeo.columbia.edu/). The NMME is a collection of coupled atmosphere–ocean–sea ice–land global seasonal forecast models developed by American and Canadian modeling centers, each with unique formulations and initialization methods. We selected the subset of models based on the availability of atmospheric variables and sufficient ensemble members (10–28 members depending on the model; Table 1). We excluded models that had fewer than six ensemble members to enable the identification of three unique high- and low-error forecasts from each model. All forecasts are interpolated to a common 1° latitude by a 1° longitude grid and comprise forecasts extending a minimum of 9.5 months from initialization at monthly resolution. For CanCM4i, GEM-NEMO, and COLA-RSMAS-CCSM4, all forecasts were initialized on the first of the month (0000 UTC). For NCEP-CFSv2, four ensemble members were initialized at 0000, 0600, 1200, and 1800 UTC every fifth day and organized into monthly starts containing the 6–7 pentad initializations centered on the first of the month. From each model, we retrieved SST, SLP, and geopotential height at 200 hPa (H200) for 1982–2020, combining hindcast and real-time forecast periods to span the 39-yr range. The real-time period represents the interval throughout which a model has been used to actively produce forecasts within the NMME, beginning between 2011 and 2019 among the four models analyzed here (Table 1). Retrospective forecasts (hindcasts) span 1982 through the start of the real-time period, produced with identical prediction systems and initialization procedures to the real-time forecasts. For each forecast, we consider lead times of 0.5–9.5 months. To minimize bias from long-term model drift, we calculate lead-time-dependent anomalies for each variable by subtracting the model ensemble mean lead-time-dependent climatology (Stock et al. 2015). The multimodel ensemble mean forecast is produced by averaging the ensemble mean forecasts across all models, such that models rather than individual ensemble members are equally weighted.

We verify forecasted SSTAs against NOAA's Optimum Interpolation Sea Surface Temperature, version 2 (OISSTv2; Table 1) (Banzon et al. 2016; Reynolds et al. 2007). OISSTv2 is available at 0.25° horizontal resolution, which we regrid to 1° resolution using bilinear interpolation for comparison to the NMME grid. We verify the atmospheric variables against monthly mean fields of the mean SLP and geopotential height at 200 hPa from the fifth major global reanalysis produced by European Centre for Medium-Range Weather Forecasting (ECMWF) (ERA5) (Table 1) (Hersbach et al. 2020). For all verification data, anomalies are calculated relative to the 1982–2020 climatology.

We further compare the skill of the NMME SST forecasts to that of persistence forecasts generated from the OISST verification anomalies, in which the SSTA of the month preceding forecast initialization is propagated for all forecast lead

TABLE 2. Outline of forecast conditioning experiments, denoting the variable, lead time, metric (absolute error or SCC), and region on which forecasts are conditioned. The skill of conditioned forecasts is examined for multiple variables.

Experiment No.	Variable	Lead time	Metric	Region
1	SST	2.5	Error	CCLME
2	SLP	1.5	SCC	15°–60°N, 170°–280°E
3	H200	1.5	SCC	25°–60°N, 180°–260°E
4	PMM	1.5	Error	−20°–60°N, 150°–260°E

times. The persistence forecast skill calculated as the anomaly correlation coefficient (ACC) is thus equivalent to the autocorrelation function of SSTA for the month prior to initialization. Because SSTA predictability partially results from the thermal inertia of the ocean, comparison of the skill of dynamical forecasts relative to that of a persistence forecast provides insight into the benefit gained from the use of dynamical models (Stock et al. 2015).

b. *Skill evaluation*

We evaluate forecast skill using ACCs and root-mean-square error (RMSE) as a function of initialization month and lead time, as in prior studies (Hervieux et al. 2019; Jacox et al. 2019; Stock et al. 2015). ACC values are considered significant for $p < 0.05$. Significant improvements in skill above persistence and between conditioned forecasts are based on Fisher's R-to-Z transformations of ACC values, which produce Z distributions with N-3 degrees of freedom that we assess with one-sided z-score tests at the confidence level of $p < 0.10$, as in Stock et al. (2015).

c. *Forecast conditioning*

Our generalized approach to calculating conditional skill is as follows: for every forecast initialization (12 months × 39 years), we rank each model's ensemble members according to performance at a short lead time (1.5 or 2.5 months), as judged by the mean absolute error or spatial correlation coefficient (SCC) calculated within a prescribed region for one field (i.e., SST, SLP, or H200 anomalies). Based on the chosen error metric, we then extract the three best and three worst ensemble members from each model for each forecast initialization and use them to construct "best" and "worst" forecast ensembles with 12 ensemble members each (3 ensemble members × 4 models). This process yields 39 composite best and worst forecast ensembles (1 yr^{-1}) for each initialization month (January–December).

Following the above steps, we perform a sequence of conditioning experiments to evaluate the relationships in forecasted fields and their skill across lead times, outlined in Table 2. We choose to condition forecasts at lead times of 2.5 and 1.5 months to balance the amount of dispersion and skill among forecasts, but note that conditioning at other lead times generates qualitatively similar results. For conditioning on the PMM (Table 2, Experiment 4), we project forecast SSTA for each member and at each time step on the PMM SSTA regression pattern (Fig. 1d) within 20°S–60°N and 150°E–120°W and calculate the absolute error of this index relative to the same calculated from OISST observations. The PMM SSTA regression pattern was produced by regressing OISST SSTA from 1982 to 2020 on the PMM SST index retrieved from the NOAA Physical Sciences Laboratory (https://psl.noaa.gov/data/timeseries/monthly/PMM/). We consider a larger region than used in the maximum covariance analysis to define the PMM in order to capture the SSTA variability associated with the mode throughout the CCLME.

Finding the best-performing forecasts among the ensemble members of each model rather than across all models retains the benefits of the multimodel approach (Hervieux et al. 2019; Hagedorn et al. 2005), while allowing for the differentiation of performance stemming from errors in initialization. The difference in skill between best and worst composite forecasts ($\Delta ACC = ACC_{Best} - ACC_{Worst}$) represents the gain in forecast skill generated by selecting forecasts based on our conditioning metric. Positive ΔACC denotes that the lowest error forecasts at the evaluated lead time (the best) outperform the highest error forecasts (the worst) at some future lead time. We denote the variable for which we calculate skill in the subscript (e.g., ΔACC_{SST}). Although the changes in the skill of the best and worst forecasts are not symmetric in magnitude when compared to the ensemble mean, similar patterns suggest that the same processes contribute to their skill changes (Fig. S1 in the online supplemental material). Because we are interested in examining these processes, we justify using the best–worst skill (i.e., ΔACC) to maximize this signal.

3. Results and discussion

a. *Skill of SSTA ensemble mean and persistence forecasts in the CCLME*

Prior studies have evaluated the performance of SSTA forecasts derived from ensemble means of single NMME models, multimodel ensembles, and persistence in the CCLME and demonstrated significant skill by each for lead times of several months (Hervieux et al. 2019; Jacox et al. 2019; Stock et al. 2015). Here, given an extended record ($n = 39$ years) and an overlapping but differing subset of the NMME, we find multimodel ensemble mean skill similar to those in earlier analyses (Figs. 2 and 3). Skill of the ensemble mean is significant for nearly all lead times and initialization months, with maximum anomaly correlation coefficients near 0.85 at the 0.5-month lead time decaying below 0.5 in 3–7 months, depending on the season of initialization (Fig. 2; Fig. S2). While the skill of the multimodel ensemble mean (Fig. 3, black dashed lines) is always higher than persistence (Fig. 3, black dotted lines) through the 9.5-month lead, rarely is the difference above persistence significant at the 90% confidence threshold as noted in previous studies (Fig. 2) (Stock et al. 2015). In the CCS, simple persistence forecasts have

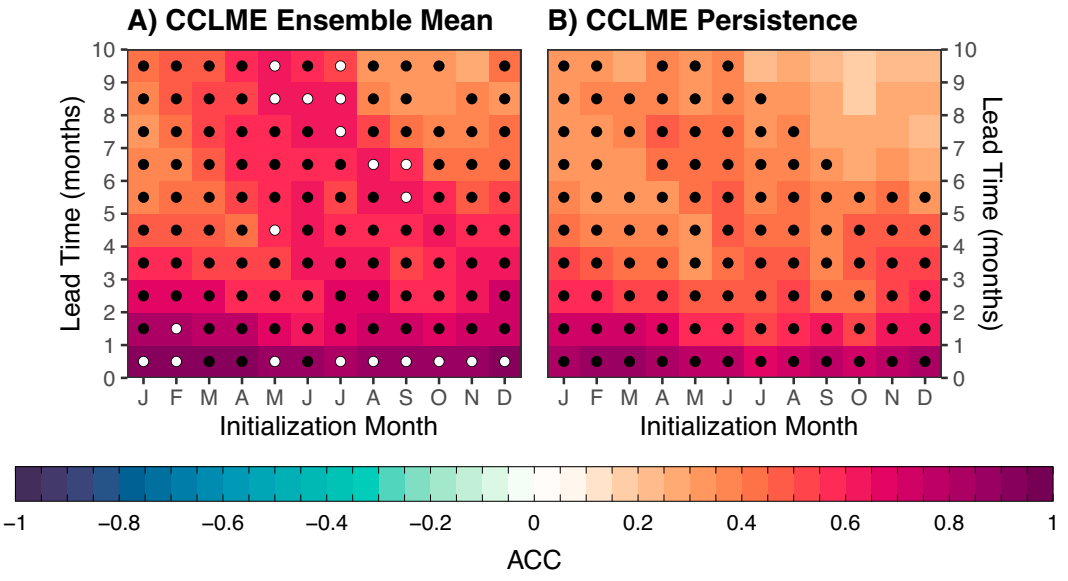

FIG. 2. Matrices showing the SSTA prediction skill in the CCLME as evaluated by ACC by initialization month and lead time, for the (left) multimodel ensemble mean and (right) persistence forecasts. Black dots indicate correlation coefficients significant at $p < 0.05$, and white dots indicate that the skill of the ensemble mean is significantly higher than the skill of persistence, as evaluated by Fisher's R-to-Z transformation at $p < 0.10$, as in Stock et al. (2015). See Fig. S2 for the same skill plotted by verification month and lead time.

lasting, significant skill (>6 months) due to the thermal inertia of the mixed layer and additional processes such as the reemergence of subsurface anomalies by deepening of the mixed layer in winter (Alexander et al. 1999).

Matrices showing ACC values for each initialization month and lead time reveal temporal patterns in both the predictability of the ocean–atmosphere system and the performance of the forecast system (Stock et al. 2015). For the ensemble mean (Fig. 2), a diagonal band of elevated values beginning in fall and winter initializations spreads to the upper-left corner, indicating high skill for forecasts of winter and early spring, regardless of when they are initialized (see Fig. S2 for skill plotted by the verification month). This swath of elevated skill in CCLME SSTA forecasts likely reflects the dominant role of ENSO in generating predictability, which provides dynamical skill and persistence concentrated in the late winter and early spring (Jacox et al. 2019; Stock et al. 2015). Patterns in RMSE are qualitatively similar (Fig. S3), and for the following analyses, we present ACC values only.

b. *Experiment 1: Forecasts conditioned on CCLME SSTA errors*

1) TEMPORAL PATTERNS IN SSTA CONDITIONAL SKILL

We first evaluate relationships between errors in predicted SSTA in the CCLME at short lead times (in this case, 2.5 months) and SSTA prediction skill in the CCLME for those same forecasts at longer lead times (Table 2, Experiment 1). If low-error forecasts (i.e., the best conditioned forecasts) correspond to significant improvements in skill at later lead times, this would indicate that successful simulation of path-dependent dynamics could increase confidence in long-lead prediction.

In the 2.5-month lead on which forecasts are conditioned, the best composites rebound to skill values similar to that of the 0.5-month lead (~0.85), whereas the worst composites decline to nonsignificant values (<0.25) (Figs. 3 and 4a–c). The rebound in skill of the best forecasts suggests that the occurrence high skill in the 2.5-month lead is random rather than persistent from the forecast initialization. Yet, the best skill continues to exceed the worst for the remainder of the forecasts for most initialization months, resulting in positive ΔACC_{SST} values (Fig. 4c). Positive ΔACC_{SST} values are significant for 3–5 months, beginning in the month preceding conditioning and continuing for up to 3–4 months following.

The best composites show significant skill for nearly all initialization months and lead times (Fig. 4a), with more instances of skill significantly above persistence relative to both the ensemble mean and the worst composites for all initialization months. Significant skill for worst composites is limited to the first two initialization months and the central swath of the initialization month-lead-time matrix (Fig. 4b). Asymmetry between the best and worst forecasts relative to the ensemble mean suggests that the ensemble mean captures some additional skill relative to the full range of forecasts produced (Fig. S1); forecasts showing the highest errors describe less probable pathways. If we instead condition on the 3.5-month lead, the pattern is comparable, shifted one lead time longer (Fig. S4).

Although the ranking of forecasts is similar among all initialization months, the along-lead-time trajectories of the

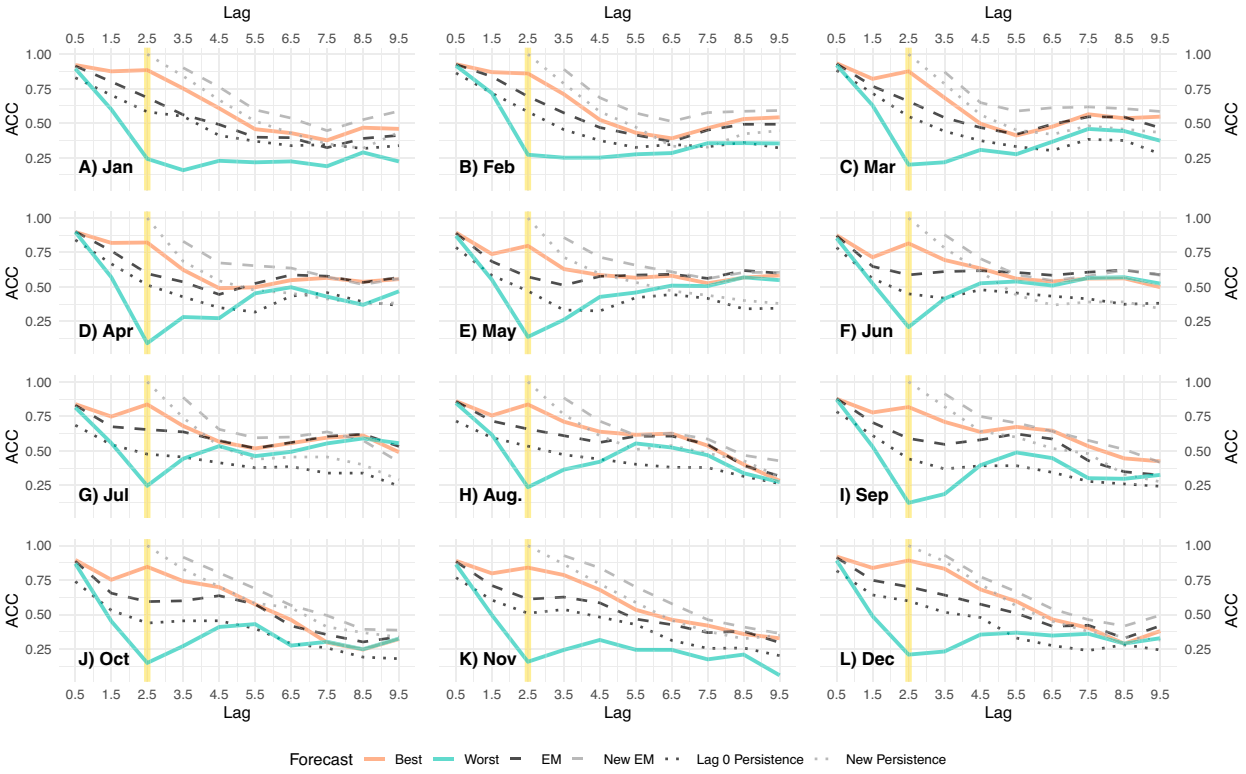

FIG. 3. Mean skill of SSTA forecasts in the CCLME conditioned on local SSTA error at lag 2.5 for each initialization month by lead time: "best" composite (orange), "worst" composite (green), full EM (black dashed), lag zero persistence (black dotted), ensemble mean of a new dynamical forecast initiated 3 months following the original initialization (gray dashed), and a persistence forecast initiated at lag 2.5 (gray dotted). The yellow vertical line denotes the lead time at which forecasts are conditioned.

evolution of skill differ (Fig. 3). Whereas the best skill declines following the conditioning lead, the worst skill either plateaus or rebounds in the following time steps, affecting the time scale at which the skill of the two composite forecasts converges. For the late winter and early spring initializations (January–April), the best and worst maintain a gap through the end of the forecast period, demonstrating that conditioning performed in the spring provides a lasting increase in skill. On the other hand, for the late spring and early summer initializations, the skill of the best and worst composites fully converges in 2–4 months, demonstrating that low errors in SSTA in the summer do not correspond to better prediction through the following fall and winter seasons. The more sustained difference in skill between the best and worst forecasts for forecasts initialized in the winter relative to forecasts initialized in the spring and summer likely reflects the influence of more persistent SSTA associated with the winter occurrence of a deeper mixed layer and processes including the seasonal footprinting mechanism (Vimont et al. 2001, 2003; Chiang and Vimont 2004).

Both the best and worst composite forecasts have significant skill along the central ridge of the initialization month-lead-time matrix, minimizing ΔACC_{SST} for these periods. That conditioning does not produce a difference in skill here suggests that long-lead ENSO-related predictability is robust to small errors in initialization or early lead times and that differentiation in skill among the best and worst composite forecasts results from a secondary driver of skill more variably represented among ensemble members (Fig. S5).

The sustained advantage of forecasts with low errors at early lead times demonstrates that better early predictions beget better later predictions of SSTA in the CCLME. From a strictly operational perspective, the magnitude of this advantage would determine its utility; if the best composite were to outperform a dynamical forecast initiated at the lead time of conditioning (Fig. 3: gray dashed lines), this would provide an opportunity to leverage the forecasts that had "learned" something about dynamics early on to produce a more confident forecast further out. We find that this is not the case (after two additional months, the best conditioned forecasts perform similarly to a newly generated persistence forecast but do not outperform the ensemble mean of a new dynamical forecast; Fig. 3: gray dotted lines). However, the fact that conditioned forecasts maintain disparities in skill provides an opportunity to evaluate why forecasts succeed or fail.

2) SPATIAL PATTERNS IN CONDITIONAL SKILL OF SSTA FORECASTS

Broadening our lens to the greater North Pacific, we observe that ΔACC_{SST} values of forecasts conditioned within

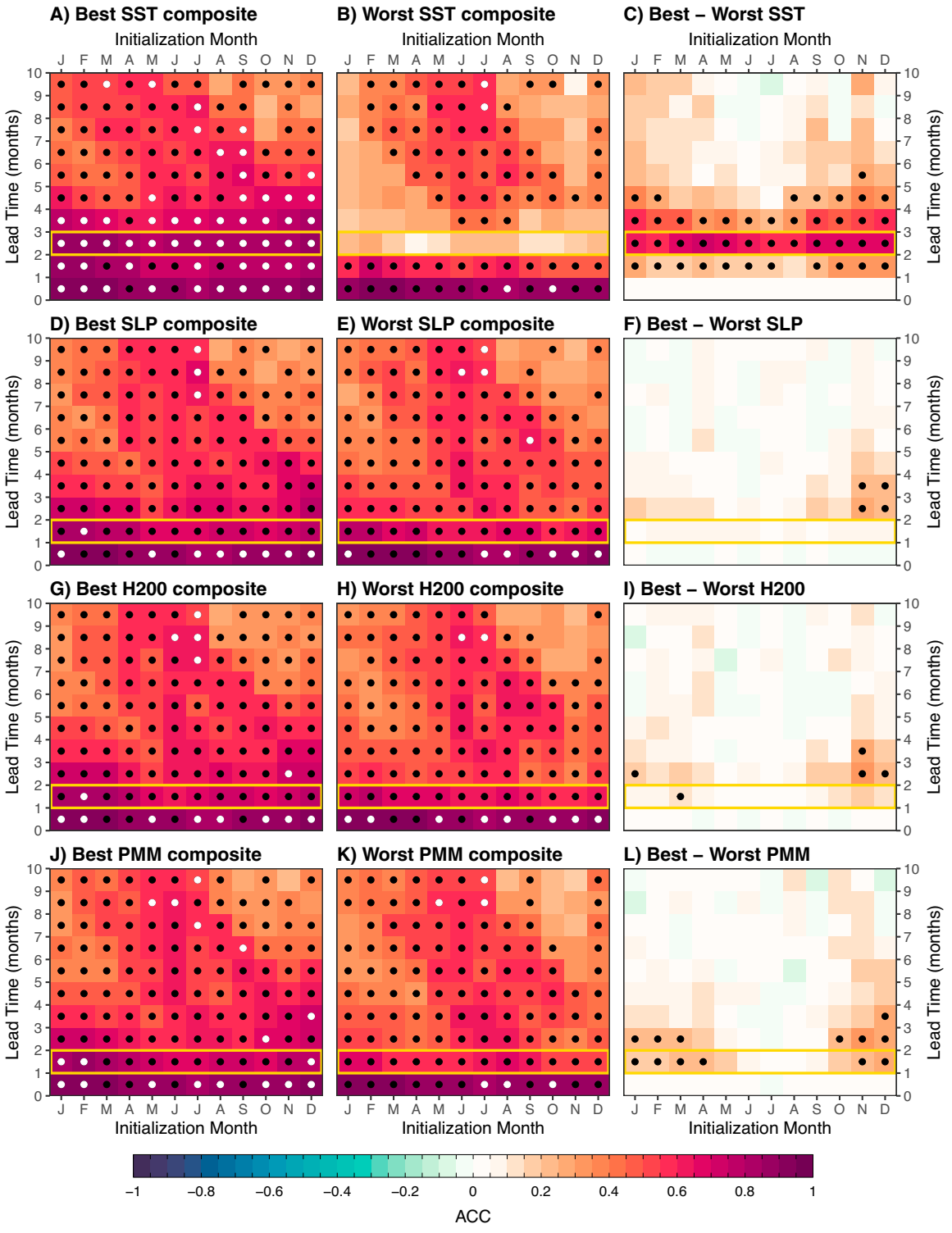

FIG. 4. Skill matrices showing SST prediction skill by initialization month and lead time in the CCLME for a series of conditioning experiments (Table 2). All plots show the mean SSTA skill in the CCLME, as conditioned on different variables. The left column shows the skill of the best composite forecasts, the center column shows the skill of the worst composite forecasts, and the right column shows the difference in the best and worst skills (ΔACC). For the left and center columns, black and white dots are as in Fig. 2. For the right column, black dots indicate that the best skill is significantly higher than the worst skill, as determined from Fisher's R-to-Z transformations at $p < 0.10$.

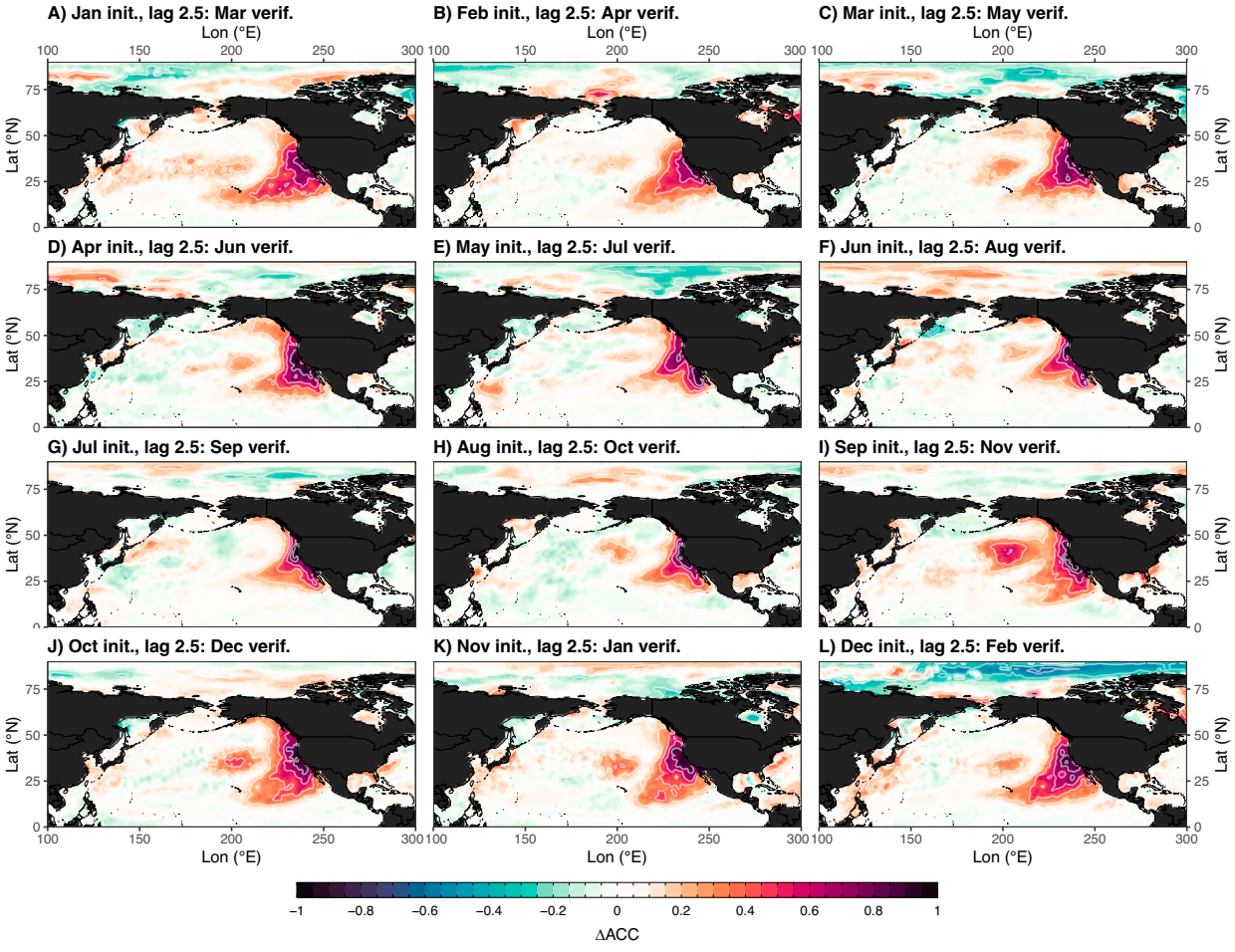

FIG. 5. Differences in the skill of composite best and worst SST forecasts (ΔACC_{SST}) at a lead time of 2.5 months for forecasts conditioned on SSTA errors within the CCLME at the 2.5 lead time for all initialization months (Table 2, Experiment 1). White contours are spaced at increments of 0.2.

the CCLME exhibit spatial footprints extending well beyond the CCLME domain (Fig. 5). For all initialization months, the most positive ΔACC_{SST} values are concentrated along the west coast of North America during the 2.5-month lead time of conditioning, as expected given that the best and worst forecasts were chosen based on this region and at this time. However, the domain of positive values extends southwest to about 15°N and 190°E in a footprint resembling the positive SSTA loadings of the PMM (Fig. 1d). A secondary patch of positive ΔACC_{SST} in the central North Pacific coincides with the negative loading of a positive PMM centered near 30°–45°N and 160°W (note that correctly forecasting negative SSTA results in positive skill). Although most pronounced at the 2.5-month lead, a pattern of positive ΔACC_{SST} values persists for a minimum of 2 months (Fig. S6) and up to 8 months when forecasts are conditioned in the late winter–early spring when the PMM is most active (Fig. S7). Spatiotemporal linkages in skill with the footprint of the PMM suggest that predictability in the CCLME could, in part, derive from the unfolding of deterministic processes associated with this mode.

3) ATMOSPHERIC PRECURSORS TO SKILLFUL SSTA PREDICTION

Thus far, we established that members of the forecast ensemble with low errors in SSTA at early lead times demonstrate sustained improvements in SSTA prediction for several months in the CCLME and find evidence that these improvements are perhaps linked to the PMM. We next ask, can other variables, such as those of the more rapidly evolving atmosphere, provide reliable precursors of the more skillful forecasts? The potential benefits of identifying such precursors are twofold: 1) to enable the early identification of forecasts more likely to be successful long term and 2) to understand the mechanisms underlying the success or failure of a given forecast. To investigate atmospheric precursors that lead to skillful SSTA prediction, we next compare the composite representation of North Pacific circulation preceding high and low errors in forecasted CCLME SSTA during the onset of the positive PMM phase. We then examine the skill of sea level pressure anomaly (SLPA) and H200a across all years regardless of the PMM state in these same forecasts.

(i) Evolution of atmospheric anomalies for a strong positive PMM progression

The PMM develops following stochastic atmospheric variability of the NPO, with SLPA leading the development of SSTA across the subtropical and northeastern Pacific by 2–3 months (Fig. 1d) (Rogers 1981; Amaya 2019; Vimont et al. 2003; Chiang and Vimont 2004). In this section, we examine the progressions of North Pacific SLPA during the onset of the positive phase of the PMM among forecasts conditioned on SSTA in the CCLME, considering SLPA averaged for approximately the upper tercile of PMM values during the month in which forecasts are conditioned (lead 2.5). We provide the February initialization as an example sensitive to the winter–spring onset of the PMM, but note that a similar progression occurs for all winter–spring initialization months (Figs. S8 and S9). For forecasts initialized in summer, only weak anomalies are observed even for the strongest PMM years, consistent with the reduction of large-scale pressure gradients in the North Pacific in this season (Fig. 1a).

The best and worst forecasts initialized in February, as judged by 2.5-month lead SSTA errors, all show the characteristic NPO SLPA dipole at the 0.5-month lead, with negative SLPA centered within 30°–45°N and 195°–225°E and positive SLPA along the Gulf of Alaska (GOA) (Fig. 6). Comparable pattern correlations with the observed SLPA among the best (0.58), worst (0.58), and ensemble mean (0.60) indicate similar fidelity among forecasts in the weeks immediately following initialization, although SLPAs are the strongest in the best forecast. The multimodel ensemble spread, as measured by the pooled standard deviation across all members, is low overall, with a maximum associated with variations in the depth and positioning of the dipole structure.

However, the forecasts diverge in the following month. By the 1.5-month lead, the best forecasts show deepening and expansion of the negative lobe of the dipole, whereas the worst forecasts show westward propagation of positive SLPA from North America throughout the GOA. Although the ensemble mean shows persistence of the SLPA dipole similar to the best forecasts, further development of low-pressure anomalies is tempered. In lead 1.5, the best forecasts have a mean pattern correlation coefficient (0.39) significantly higher than the ensemble mean (0.29) and the worst (0.19). The center of action of the ensemble pooled standard deviation intensifies relative to the prior lead, remaining centered over the Aleutian Islands. Lead 1.5 directly precedes our conditioning on lead 2.5 SSTA errors, and the atmospheric signals during this time step represent those driving SST variations in the subsequent month.

By lead 2.5, SLPAs in all forecast composites greatly diminish, as do the pattern correlation values. This reduction in the magnitude of anomalies could, in part, reflect springtime weakening of the AL but is likely exacerbated by signal loss driven by the stochastic dispersion of the ensemble members within each forecast. Such signal loss, consistent with a greater reduction of anomalies within the forecasts than observations (Fig. 6c versus Figs. 6f,i,l), would imply that the atmospheric state in the month of conditioning is less important to SSTA than the atmospheric state in the preceding month. Indeed, worse performance in the SLPA prediction of the best composite in the lead time of conditioning on SSTA (and those following, not shown) suggests that the predominant signal is that of the atmosphere imprinting on the ocean rather than vice versa. The pooled ensemble standard deviation reaches a maximum during the 1.5-month lead time shown, suggesting that ensemble members diverge substantially in the positioning of pressure systems, leading to high standard deviation in regions of strong SLP gradients within only a month. Reduced standard deviation in the lead 2.5-month SLPA forecast of April likely is associated with the seasonal relaxation of SLPA gradients in the Northern Hemisphere through springtime.

Despite the high standard deviation associated with the positioning of pressure systems, the progressive development of widespread anomalous low pressure in the ensemble mean through the 0.5- and 1.5-month leads indicates a deterministic component in the evolution of SLPA from initialization (although the SLPA at initialization could be random). Moreover, the signal among the worst composite forecasts indicates that forecasts are consistently producing high errors in SSTA in the CCLME for the same reasons—in this case, specifically, the development of anomalous high pressure over the GOA. The better-performing forecasts more accurately simulate relevant regional atmospheric dynamics: by having correct atmospheric circulation, forecasts are more likely to accurately represent the strength and location of surface winds and fluxes, giving rise to more accurate SSTA.

(ii) North Pacific SLP precursors to CCLME SSTA forecast skill

We next examine the difference in the skill of SLPA for best and worst forecasts, conditioned on SSTA in the CCLME at the 2.5-month lead. In the previous section, we found that SSTAs were most influenced by SLPA 1 month prior, so here we consider the SLPA forecast skill at the 1.5-month lead directly preceding the 2.5-month lead conditional skill assessment for SSTA.

Spatial patterns of ΔACC_{SLP} (i.e., the best minus worst skill difference for SLPA forecasts) in the 1.5-month lead are shown in Fig. 7. A hotspot of positive ΔACC_{SLP} adjacent to the CCLME is present for all initialization months, although the location and extent vary seasonally. For most initialization months, the region of positive values is centered within 30°–45°N and 210°–225°E, elongated in the east–west direction, directly west of the CCLME. For some initialization months (November–January), the area spans the North Pacific midlatitudes fully. For approximately half of the initialization months, a second node of positive ΔACC_{SLP} emerges over western North America. For the late spring and summer initializations, the region of positive ΔACC_{SLP} retracts to a small center of action just off the U.S. West Coast.

Areas of strongly positive ΔACC_{SLP} align with the low-pressure anomalies associated with the NPO (see the previous section), which modulate the northerly winds over the CCLME along the eastern edge of the cyclonic circulation. The region of elevated skill centered near 35°N and 220°E sits between the

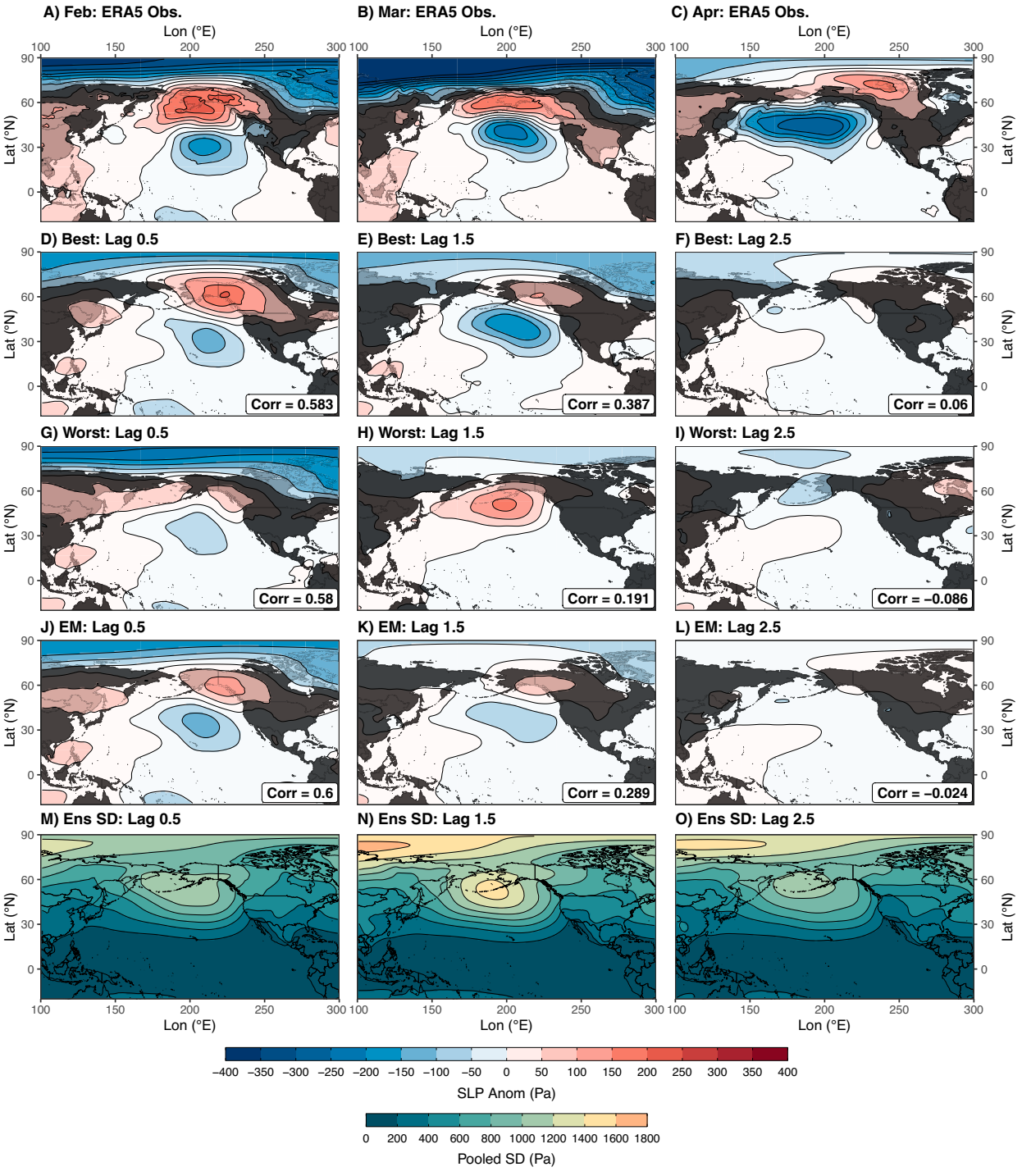

FIG. 6. Mean monthly progression of SLP anomalies during strongly positive PMM years (n = 15) for forecasts initialized in February (see Figs. S8 and S9 for other initialization months). (a)–(c) Observations from the ERA5 reanalysis. (d)–(f) The progression of SLP in the best composite forecast (based on SST anomalies in the CCLME at lag 2.5). (g)–(i) The progression for the worst composite forecast. (j)–(l) The multimodel EM. (m)–(o) The pooled standard deviation for ensemble members of the multimodel ensemble. Mean pattern correlation coefficients (uncentered) of forecasts and observations within the box 15°55°N and 180°–250°E are shown in the lower right-hand corner.

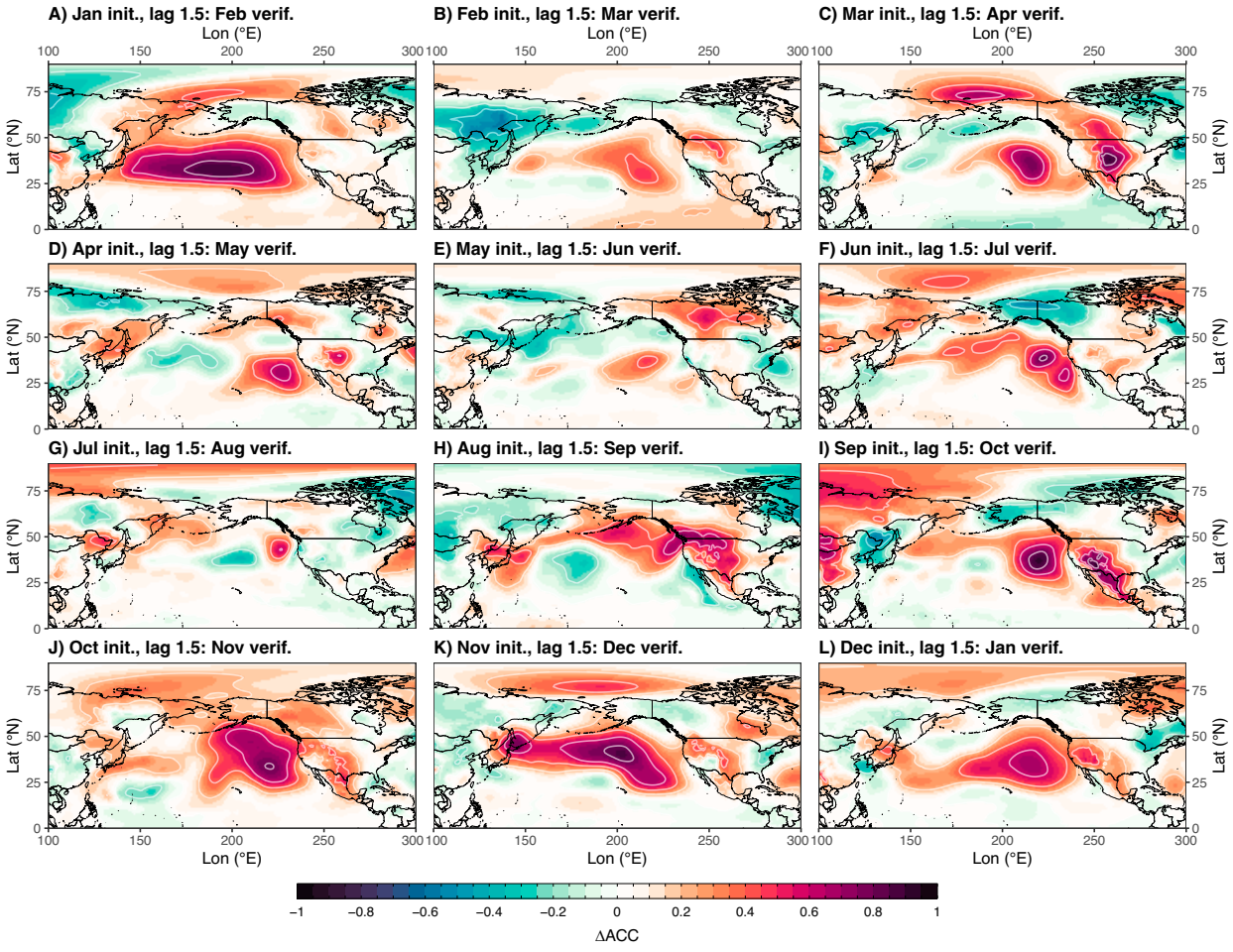

FIG. 7. Maps of ΔACC_{SLP} showing differences in the skill of composite best and worst SLPA forecasts for all initialization months at a lead time of 1.5 months, the lead time prior to forecast conditioning on SST in the CCLME at lead 2.5 (Table 2, Experiment 1). White contours are spaced at increments of 0.2.

climatological mean positions of the winter AL and North Pacific High, indicating the dependence of SSTA skill on accurately simulating the strength and positioning of the southern lobe of the NPO. The secondary node of positive ΔACC_{SLP} over North America corresponds to continental high pressure, which further modulates nearshore winds. Even in the absence of strong signals in the summer, the small regions of positive ΔACC_{SLP} indicate that local anomalies likely influence winds along the eastern Pacific margin at reduced spatial scales, continuing to provide skill in the CCLME during seasons less sensitive to anomalous basin-scale circulation. Similar seasonal patterns are seen in ΔACC_{H200} conditioned on SSTA errors in the CCLME (Fig. S10), with centers of action resembling those of the west Pacific (WP) teleconnection, taken to be the upper-atmosphere expression of the NPO (Linkin and Nigam 2008; Wallace and Gutzler 1981).

c. *Experiments 2 and 3: Forecasts conditioned on atmospheric fields*

Anomalous SLP and surface winds typically lead the development of PMM SSTA by 1–2 months (Chiang and Vimont 2004). Identifying the atmospheric expression of the onset of coupled ocean–atmosphere variability therefore provides an opportunity to enhance SSTA prediction on this time scale. Following the establishment of regions of high skill in SLPA prediction in the months preceding low errors in SSTA forecasts (previous section), we invert the analysis to condition forecasts on SLP and H200 anomalies during the 1.5-month lead in regions demonstrated to be sensitive to forecast performance in the CCLME (Table 2, Experiments 2 and 3).

Positive ΔACC_{SST} values are observed for winter and spring initializations in the months directly following conditioning on 1.5-month lead SLPA (Figs. 4d–f). Significant increases in the skill of the best over the worst are limited to November and December initializations for the 2 months after conditioning, corresponding to forecasts of January–March. The similarity of results from conditioning on SLPA and H200a demonstrates that the patterns in the SSTA skill are robust to the specific variable used (Figs. 4g–i).

When conditioning best and worst forecasts on SLPA at lead 1.5, positive ΔACC_{SST} values at lead 2.5 occur in broad regions of the northeast Pacific for all initialization months,

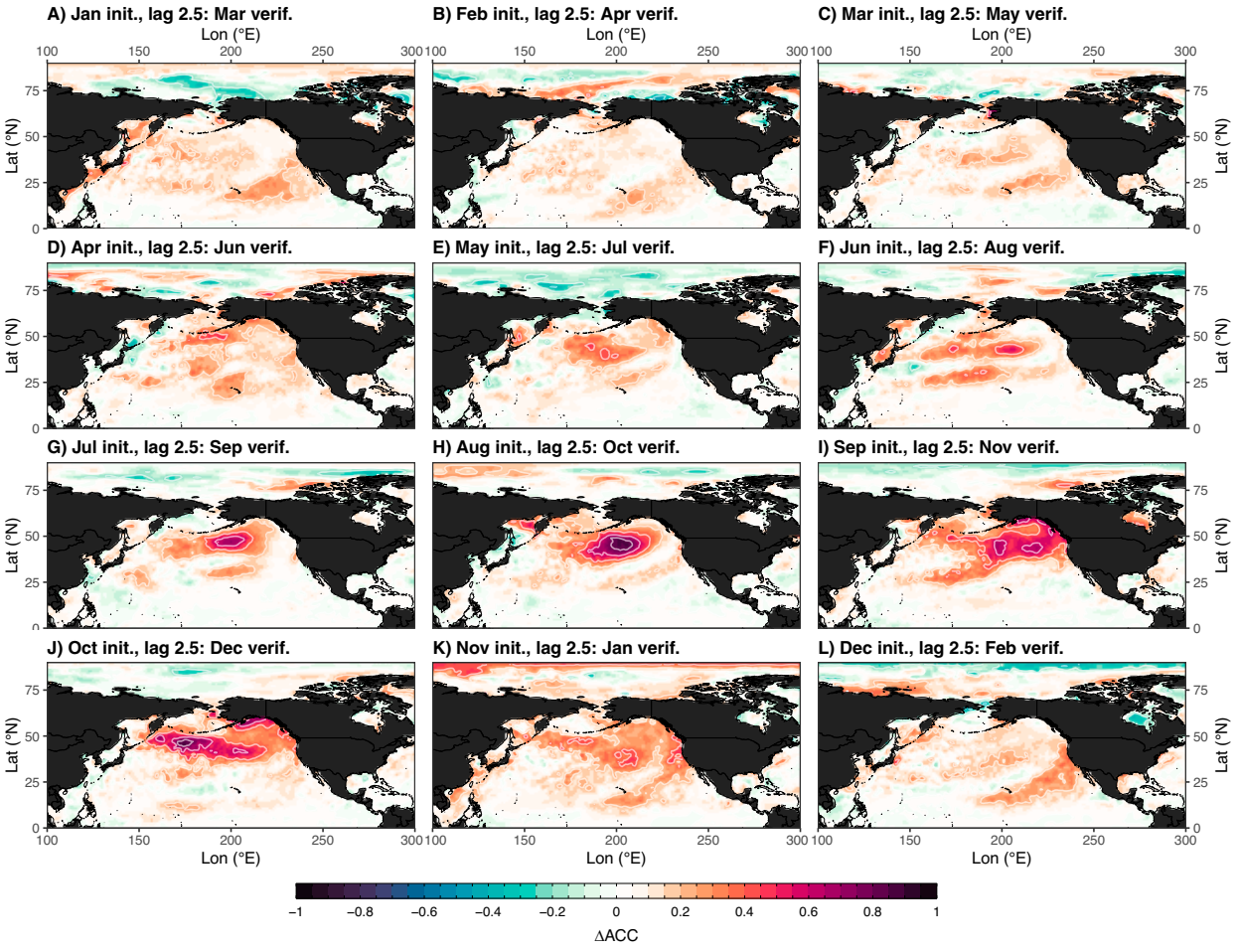

FIG. 8. Maps of ΔACC_{SST} showing differences in the skill of composite best and worst SSTA forecasts for all initialization months at a lead time of 2.5 months, the lead time following conditioning on SLP at lead 1.5 (Table 2, Experiment 2). White contours are spaced at increments of 0.2.

though the location and magnitude of maxima vary (Fig. 8; for conditioning on H200a, see Fig. S11). For the late fall through early spring initializations (November–March), aspects of the PMM footprint are present, most notably as a tongue extending southwestward from the CCLME toward the tropics where SSTA loads positively on the PMM. However, the PMM-like pattern breaks down in the spring, as the maximum ΔACC_{SST} values detach from the eastern Pacific coast and migrate into the north-central Pacific before concentrating northward along the Aleutian Islands and throughout the GOA in the summer. Positive ΔACC_{SST} values then extend toward the midlatitudes in August–October initializations, returning to the PMM-like pattern in the late fall. Strong signals are present for most initialization months but are not always evidently linked to the PMM, demonstrating the importance of seasonality in mechanisms underlying predictability.

While most initialization months show positive ΔACC_{SST} throughout the CCLME at the 2.5-month lead (Fig. 4f), the magnitude of the signal varies throughout the LME and tends to be strongest in the north (Fig. 8). Prior studies have noted a similar gradient in the skill of SSTA prediction in the CCS, consistent with the spatial patterns of ENSO influence on wind-driven upwelling (Jacox et al. 2019). While winds south of Cape Mendocino (~40°N) are persistently equatorward and upwelling favorable, winds to the north vary seasonally depending on the positioning of the AL and, therefore, can promote upwelling or downwelling depending on the configuration of SLPA. As a result, the same shift in the AL would drive more dramatic consequences in the north than the south, and the associated anomalies in winds and SLP in the northern sector would provide more leverage in determining the SSTA skill. In the central and southern sectors, anomalous SLP and winds still alter SSTA but provide more modest increases in the SSTA skill; these latitudes are more strongly driven by remote ocean teleconnections than by atmospheric forcing (Frischknecht et al. 2015).

d. Experiment 4: Forecasts conditioned on the PMM

Across our analyses, the PMM presents as a throughline of the coupled ocean–atmosphere variability associated with

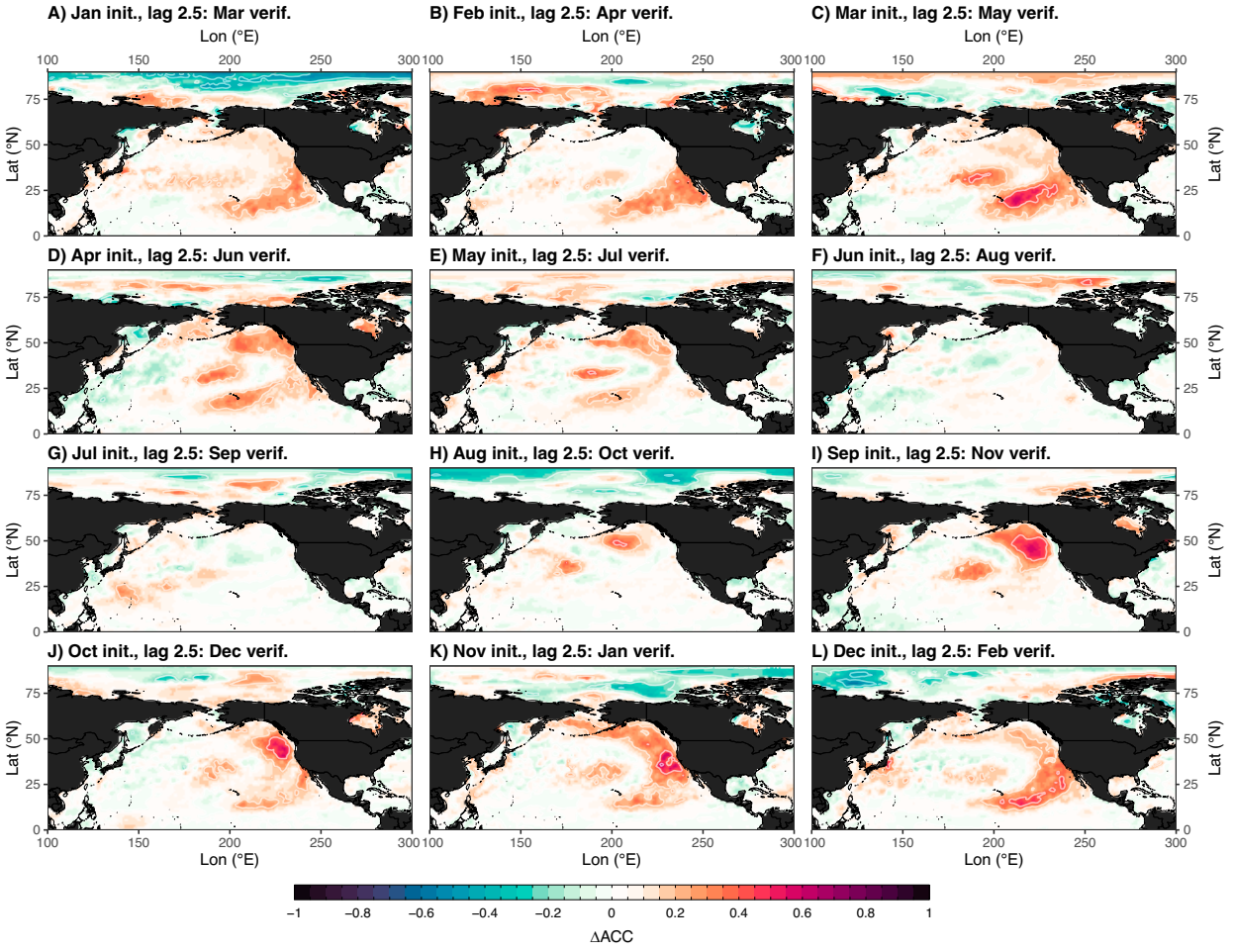

FIG. 9. Maps of ΔACC_{SST} showing differences in the skill of composite best and worst SSTA forecasts for all initialization months at a lead time of 2.5 months, the lead time following conditioning on PMM at lead 1.5 (Table 2, Experiment 4). White contours are spaced at increments of 0.2.

prediction skill in the CCLME and the northeast Pacific. As seen in the spatial pattern of SSTA regressed on the PMM index (Fig. 1d), SSTA variability in the CCLME is linked to the PMM and so it follows that predictability in the CCLME arises in association with this mode or from a shared pattern of atmospheric forcing. We finally consider the increases in the skill of SSTA forecasts resulting from conditioning on the error of the PMM index at the 1.5-month lead (Table 2, Experiment 4). Lacking fields of forecasted winds, we condition forecasts on the SSTA pattern of the PMM, which is an imprint of the coupled mode but may not enable the detection of signals specifically driven by leading atmospheric anomalies.

As among the forecasts conditioned on SLP and H200 anomalies, positive ΔACC_{SST} values produced by conditioning on the PMM are confined to fall and winter initialization months (Figs. 4j–l). For October–April initializations, positive ΔACC_{SST} values occur for 1–7 months following conditioning, with significant increases of the best over the worst for 1–2 months in November and December. Conditioning on the PMM provides no advantage in skill for forecasts initialized in June through September.

Both the CCLME and the broader North Pacific exhibit sensitivity of SSTA forecast skill during and following the time step of conditioning on the PMM, whereas forecasts of SLPA and H200a conditioned on the PMM show strong signals in the lead time preceding conditioning (Figs. S12 and S13). These patterns in timing are overall consistent with the canonical timing of the PMM, in which the variance of expansion coefficients peaks in the late winter/early spring for wind and in spring for SST and the seasonal footprinting mechanism (Chiang and Vimont 2004; Amaya 2019).

Positive ΔACC_{SST} values produced by conditioning on the PMM show characteristic aspects of the mode's SSTA for fall through spring initialization months (November–May) (Fig. 9). The tongue extending southwestward from the south-central CCLME is prominent in the spring, whereas the full horseshoe shape extends along the northeast coastal Pacific into the GOA in the fall through early winter, with a patch of positive values that coincides with the negative loadings near

35°N, 190°E. As in the atmospheric conditioning (Table 2, Experiments 2 and 3), a strong signal in the GOA emerges for August through November initializations.

Initiation of the PMM has been traced back to random perturbations of the NPO in winter, with those perturbations that weaken the subtropical northeasterly trade winds triggering feedbacks that develop SSTA in the north tropical Pacific and extratropics, maximizing in spring (Chiang and Vimont 2004; Xie and Philander 1994). These North Pacific SLPAs reduce northerly winds along the CCLME, reducing upwelling and evaporation and increasing downward latent heat fluxes to promote warm SSTA in a crescent-shaped footprint resembling PMM SSTA loadings (Johnstone and Mantua 2014). Linear prediction based on SLPA and the persistence of SSTA in the northeast Pacific can reproduce the pattern (closely resembling that in ΔACC_{SST} for November–January initializations) even without teleconnection to the subtropics or dynamic ocean processes (Di Lorenzo and Mantua 2016). Nonetheless, forecasts which, perhaps by happenstance, capture the early initiation of the PMM SSTA pattern or the NPO trigger and proceed to develop along a predictable pattern of SLPA and SSTA for 1–2 months enhance predictability in the CCLME in winter and spring seasons when the pattern is active. Targeting regions and periods sensitive to the onset of coupled variability could lengthen time scales of predictability sensitive to stochastically initiated wind-driven processes.

4. Summary and implications for forecasting in the CCLME and beyond

We return to the questions posed in the introduction to summarize the findings and discuss implications of the analyses performed in this study for forecasting SSTA in the CCLME, North Pacific, and elsewhere.

a. Do forecasts with low errors in SSTA in the CCLME at short lead times sustain an advantage in prediction skill at longer lead times?

We first demonstrate that forecasts that better simulate SSTA in the CCLME at short lead times retain an increase in skill above forecasts with high errors, outperforming ensemble mean and persistence forecasts for several lead times. The time scale for which the "best" forecasts outperform the ensemble mean and "worst" forecasts depend on the initialization month, with the longest-lasting increase in skill observed for forecasts initialized in the fall and winter. However, the best forecasts conditioned on SSTA in the CCLME at a 2.5-month lead do not outperform newly initialized dynamical forecasts, indicating that a new ensemble mean is likely to be more skillful than even the best subset of forecasts from an earlier initialization. The difference in the skill of SSTA prediction between the best- and worst-performing forecasts of SSTA in the CCLME shows footprints of increased skill extending beyond the CCLME, with a seasonally varying pattern resembling the PMM in the fall through spring.

b. Are there atmospheric and/or oceanic precursors associated with skillful SSTA forecasts in the CCLME?

Elevated skill is present in atmospheric variables over large regions of the North Pacific preceding low errors in CCLME SSTA, and the best SSTA forecasts are preceded by more accurate progressions of SLPA. It is well established that SSTA variability in the northeast Pacific is sensitive to large-scale atmospheric circulation (Di Lorenzo and Mantua 2016; Chhak et al. 2009; Johnstone and Mantua 2014; Capotondi et al. 2019), and most low-frequency variability in upwelling in the coastal eastern Pacific is linked to large-scale climate patterns (Di Lorenzo et al. 2013; Jacox et al. 2014). In addition to exerting a direct control on North Pacific SSTA, the atmosphere has much shorter decorrelation time scales than the ocean, resulting in earlier and larger errors in SLPA than SSTA. Identifying the onset of patterns driving predictability, which may be first detectable among atmospheric variables, may inform improvements in forecast skill in the CCLME and coastal northeast Pacific during the fall through winter. As atmospheric stochasticity presents a challenge for prediction and could limit predictability for SSTA that are sensitive to surface wind interactions, atmospheric precursors of ocean forecast skill might be most helpful on time scales of 1–2 months (i.e., for subseasonal prediction rather than seasonal). Additionally, the consequence of stochastic atmospheric forcing suggests the importance of using large ensembles and probabilistic measures of forecast skill for predicting SSTA.

This study underscores the suggestion by Stock et al. (2015) that poor performance in atmospheric models impacts ocean performance, which could be incorrectly attributed to poorly resolved regional ocean dynamics. We do not evaluate winds directly in this study due to the limited availability of output, underscoring how increased access to additional variables among NMME models (e.g., 10-m winds and wind stresses) would be useful for further forecast verification and diagnostics.

Conditioning on atmospheric anomalies shows the most promise of being operationally useful in the GOA, where signals are more pronounced than in the CCLME (Fig. 8; Fig. S11). The signal in the GOA is among the strongest observed, indicating high sensitivity of SSTA in this region to North Pacific atmospheric circulation as identified in prior studies (Di Lorenzo and Mantua 2016; Capotondi et al. 2019), particularly for summer initialization and fall verification months in our analyses. Prior studies have demonstrated skill in SSTA prediction in the GOA significantly above persistence for 6-month-lead forecasts initialized in August (Stock et al. 2015; Wen et al. 2012). Stock et al. (2015) hypothesize that this sensitivity of winter SSTA to conditions in the prior summer results from the advection of offshore SSTA and winter heat flux anomalies, with the delivery of anomalously warm, moist air by strengthened southerlies, as occurs during the positive phases of the PDO and El Niño events (Alexander et al. 2002; Mantua et al. 1997). The patterns observed here resemble coupled SLP and SST patterns leading marine heat waves in the GOA, tentatively arising from PMM–ENSO interactions and promoting the persistence of SSTA through

favorable extratropical atmospheric circulation on seasonal-to-interannual time scales (Capotondi et al. 2022; Di Lorenzo and Mantua 2016; Johnstone and Mantua 2014; Baxter and Nigam 2015; Xu et al. 2021).

c. How do atmospheric processes and coupled modes contribute to spatiotemporal patterns in predictability and prediction skill of SSTA in the CCLME and North Pacific?

Deterministic processes, such as those driving coupled ocean–atmosphere modes, are an important source of predictability and resulting prediction skill in the CCLME and elsewhere (Amaya et al. 2022; Stan et al. 2017). The PMM emerges as an additional relevant mode of coupled ocean–atmosphere variability, showing up as the characteristic spatial pattern separating the best and worst forecasts at seasonal time scales. Whereas the downstream predictability of ENSO variability in the CCLME is well represented across most forecasts, the simulation of seasonal PMM-related variability could be less robust due to its reliance on largely stochastic atmospheric triggers at the seasonal scale, contributing to greater differentiation in CCLME forecast skill among ensemble members. Variability in anomalous large-scale atmospheric circulation associated with the PMM may be secondary to the leading predictive power of ENSO, differentiating those forecasts that succeed or fail at simulating SSTA in the CCLME.

The PMM provides a link between errors in the extratropical circulation and errors in SSTA in the CCLME. Although the initiation of the PMM may be largely stochastic, the phases of growth and persistence as the mode matures provide a deterministic link between atmospheric and oceanic forecast skill. Ensemble members that best capture the PMM pattern retain an increase in skill in SSTA for several months for spring and fall initializations. Although we have focused on atmospheric controls on SSTA in the CCLME, regional ocean dynamics could additionally contribute to the persistence of the SSTA in the CCLME associated with the PMM or the NPO (Amaya 2019). Identifying the initiation of other such deterministic patterns in the preceding atmospheric state allows for the earlier detection of a pattern with enhanced predictive skill and suggests the inclusion of sea surface height (SSH) as a predictor to include the representation of coastally trapped waves, albeit at coarse resolution.

Several studies have explored the PMM as a possible precursor to El Niño (particularly Central Pacific events), with interactions between the PMM and ENSO possibly exciting each other (Larson and Kirtman 2014; Stuecker 2018; Ma et al. 2017; Amaya 2019). Long-term interactions with ENSO are thought to cause the long persistence of the SSTA of the PMM relative to its analogous mode in the Atlantic (Chiang and Vimont 2004). The longest-lasting increases in skill for forecasts conditioned on the PMM occur for October–December initializations, and persistence in the SSTA footprint of the PMM could result from interactions with ENSO, as well as from greater thermal inertia owing to a deeper mixed layer in the winter. Further understanding linkages in predictability of the PMM could support improved long-range prediction for both ENSO and SSTA in the CCLME and northeast Pacific. Further investigation of whether linked low-frequency variability in the PMM and central Pacific ENSO associated with tropical–extratropical interactions could help precondition forecasts at seasonal time scales is an avenue for future research (Capotondi et al. 2023; Stuecker 2018; Di Lorenzo et al. 2015). While tropical sources of extratropical forecast errors have received considerable attention, our study shows that extratropical atmospheric forecast errors play a considerable role in driving Northeast Pacific ocean forecast skill. These results offer North Pacific atmospheric circulation as a target for forecast model improvement that would enhance ocean forecasts as well.

Acknowledgments. This work was supported by the NOAA Climate Program Office Modeling, Analysis, Predictions, and Projections Program. We thank Antonietta Capotondi for providing helpful comments on the manuscript.

Data availability statement. All NMME forecasts used in this study are openly available from the IRI/LEDO Climate Data Library (https://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/) (Kirtman et al. 2014). OISST and ERA5 data used for forecast verification are freely available from the NOAA Physical Sciences Laboratory (https://psl.noaa.gov/) and the Copernicus Climate Change Service Climate Data Store (https://cds.climate.copernicus.eu), respectively (Banzon et al. 2016; Hersbach et al. 2020).

REFERENCES

Alexander, M. A., C. Deser, and M. S. Timlin, 1999: The reemergence of SST anomalies in the North Pacific Ocean. *J. Climate*, **12**, 2419–2433, https://doi.org/10.1175/1520-0442(1999)012<2419:TROSAI>2.0.CO;2.

——, I. Bladé, M. Newman, J. R. Lanzante, N.-C. Lau, and J. D. Scott, 2002: The atmospheric bridge: The influence of ENSO teleconnections on air–sea interaction over the global oceans. *J. Climate*, **15**, 2205–2231, https://doi.org/10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2.

Amaya, D. J., 2019: The Pacific meridional mode and ENSO: A review. *Curr. Climate Change Rep.*, **5**, 296–307, https://doi.org/10.1007/s40641-019-00142-x.

——, M. G. Jacox, J. Dias, M. A. Alexander, K. B. Karnauskas, J. D. Scott, and M. Gehne, 2022: Subseasonal-to-seasonal forecast skill in the California Current System and its connection to coastal kelvin waves. *J. Geophys. Res. Oceans*, **127**, e2021JC017892, https://doi.org/10.1029/2021JC017892.

Banzon, V., T. M. Smith, T. M. Chin, C. Liu, and W. Hankins, 2016: A long-term record of blended satellite and in situ sea-surface temperature for climate monitoring, modeling and environmental studies. *Earth Syst. Sci. Data*, **8**, 165–176, https://doi.org/10.5194/essd-8-165-2016.

Baxter, S., and S. Nigam, 2015: Key role of the North Pacific Oscillation–West Pacific pattern in generating the extreme 2013/14 North American winter. *J. Climate*, **28**, 8109–8117, https://doi.org/10.1175/JCLI-D-14-00726.1.

Becker, E. J., B. P. Kirtman, M. L'Heureux, Á. G. Muñoz, and K. Pegion, 2022: A decade of the North American Multimodel Ensemble (NMME): Research, application, and future

directions. *Bull. Amer. Meteor. Soc.*, **103**, E973–E995, https://doi.org/10.1175/BAMS-D-20-0327.1.

Brady, R. X., N. S. Lovenduski, S. G. Yeager, M. C. Long, and K. Lindsay, 2020: Skillful multiyear predictions of ocean acidification in the California Current System. *Nat. Commun.*, **11**, 2166, https://doi.org/10.1038/s41467-020-15722-x.

Capotondi, A., P. D. Sardeshmukh, E. Di Lorenzo, A. C. Subramanian, and A. J. Miller, 2019: Predictability of US West Coast Ocean temperatures is not solely due to ENSO. *Sci. Rep.*, **9**, 10993, https://doi.org/10.1038/s41598-019-47400-4.

——, M. Newman, T. Xu, and E. Di Lorenzo, 2022: An optimal precursor of Northeast Pacific marine heatwaves and central Pacific El Niño events. *Geophys. Res. Lett.*, **49**, e2021GL097350, https://doi.org/10.1029/2021GL097350.

——, and Coauthors, 2023: Mechanisms of tropical Pacific decadal variability. *Nat. Rev. Earth Environ.*, **4**, 754–769, https://doi.org/10.1038/s43017-023-00486-x.

Checkley, D. M., Jr., and J. A. Barth, 2009: Patterns and processes in the California Current System. *Prog. Oceanogr.*, **83**, 49–64, https://doi.org/10.1016/j.pocean.2009.07.028.

Chhak, K. C., E. Di Lorenzo, N. Schneider, and P. F. Cummins, 2009: Forcing of low-frequency ocean variability in the Northeast Pacific. *J. Climate*, **22**, 1255–1276, https://doi.org/10.1175/2008JCLI2639.1.

Chiang, J. C. H., and D. J. Vimont, 2004: Analogous Pacific and Atlantic Meridional modes of tropical atmosphere–ocean variability. *J. Climate*, **17**, 4143–4158, https://doi.org/10.1175/JCLI4953.1.

Chikamoto, M. O., A. Timmermann, Y. Chikamoto, H. Tokinaga, and N. Harada, 2015: Mechanisms and predictability of multiyear ecosystem variability in the North Pacific. *Global Biogeochem. Cycles*, **29**, 2001–2019, https://doi.org/10.1002/2015GB005096.

Davis, R. E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, **6**, 249–266, https://doi.org/10.1175/1520-0485(1976)006<0249:POSSTA>2.0.CO;2.

Deser, C., M. A. Alexander, S.-P. Xie, and A. S. Phillips, 2010: Sea surface temperature variability: Patterns and mechanisms. *Annu. Rev. Mar. Sci.*, **2**, 115–143, https://doi.org/10.1146/annurev-marine-120408-151453.

Di Lorenzo, E., and N. Mantua, 2016: Multi-year persistence of the 2014/15 North Pacific marine heatwave. *Nat. Climate Change*, **6**, 1042–1047, https://doi.org/10.1038/nclimate3082.

——, and Coauthors, 2008: North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophys. Res. Lett.*, **35**, L08607, https://doi.org/10.1029/2007GL032838.

——, and Coauthors, 2013: Synthesis of Pacific Ocean climate and ecosystem dynamics. *Oceanography*, **26** (4), 68–81, https://doi.org/10.5670/oceanog.2013.76.

——, G. Liguori, N. Schneider, J. C. Furtado, B. T. Anderson, and M. A. Alexander, 2015: ENSO and meridional modes: A null hypothesis for Pacific climate variability. *Geophys. Res. Lett.*, **42**, 9440–9448, https://doi.org/10.1002/2015GL066281.

Dias, J., and G. N. Kiladis, 2019: The influence of tropical forecast errors on higher latitude predictions. *Geophys. Res. Lett.*, **46**, 4450–4459, https://doi.org/10.1029/2019GL082812.

Doblas-Reyes, F. J., J. García-Serrano, F. Lienert, A. P. Biescas, and L. R. L. Rodrigues, 2013: Seasonal climate predictability and forecasting: Status and prospects. *Wiley Interdiscip. Rev.: Climate Change*, **4**, 245–268, https://doi.org/10.1002/wcc.217.

Eveson, J. P., A. J. Hobday, J. R. Hartog, C. M. Spillman, and K. M. Rough, 2015: Seasonal forecasting of tuna habitat in the Great Australian Bight. *Fish. Res.*, **170**, 39–49, https://doi.org/10.1016/j.fishres.2015.05.008.

Frankignoul, C., and K. Hasselmann, 1977: Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability. *Tellus*, **29A**, 289–305, https://doi.org/10.3402/tellusa.v29i4.11362.

Frischknecht, M., M. Münnich, and N. Gruber, 2015: Remote versus local influence of ENSO on the California Current System. *J. Geophys. Res. Oceans*, **120**, 1353–1374, https://doi.org/10.1002/2014JC010531.

Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus*, **57A**, 219–233, https://doi.org/10.3402/tellusa.v57i3.14657.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hervieux, G., M. A. Alexander, C. A. Stock, M. G. Jacox, K. Pegion, E. Becker, F. Castruccio, and D. Tommasi, 2019: More reliable coastal SST forecasts from the North American multimodel ensemble. *Climate Dyn.*, **53**, 7153–7168, https://doi.org/10.1007/s00382-017-3652-7.

Hobday, A. J., S. M. Maxwell, J. Forgie, and J. McDonald, 2013: Dynamic ocean management: Integrating scientific and technological capacity with law, policy, and management. *Stanford Environ. Law J.*, **33**, 125–165.

——, C. M. Spillman, J. P. Eveson, J. R. Hartog, X. Zhang, and S. Brodie, 2018: A framework for combining seasonal forecasts and climate projections to aid risk management for fisheries and aquaculture. *Front. Mar. Sci.*, **5**, 137, https://doi.org/10.3389/fmars.2018.00137.

Infanti, J. M., and B. P. Kirtman, 2016: Prediction and predictability of land and atmosphere initialized CCSM4 climate forecasts over North America. *J. Geophys. Res. Atmos.*, **121**, 12 690–12 701, https://doi.org/10.1002/2016JD024932.

Jacox, M. G., A. M. Moore, C. A. Edwards, and J. Fiechter, 2014: Spatially resolved upwelling in the California Current System and its connections to climate variability. *Geophys. Res. Lett.*, **41**, 3189–3196, https://doi.org/10.1002/2014GL059589.

——, C. A. Edwards, E. L. Hazen, and S. J. Bograd, 2018: Coastal upwelling revisited: Ekman, Bakun, and improved upwelling indices for the U.S. West Coast. *J. Geophys. Res. Oceans*, **123**, 7332–7350, https://doi.org/10.1029/2018JC014187.

——, M. A. Alexander, C. A. Stock, and G. Hervieux, 2019: On the skill of seasonal sea surface temperature forecasts in the California Current System and its connection to ENSO variability. *Climate Dyn.*, **53**, 7519–7533, https://doi.org/10.1007/s00382-017-3608-y.

——, and Coauthors, 2020: Seasonal-to-interannual prediction of North American coastal marine ecosystems: Forecast methods, mechanisms of predictability, and priority developments. *Prog. Oceanogr.*, **183**, 102307, https://doi.org/10.1016/j.pocean.2020.102307.

Johnstone, J. A., and N. J. Mantua, 2014: Atmospheric controls on Northeast Pacific temperature variability and change, 1900–2012. *Proc. Natl. Acad. Sci. USA*, **111**, 14 360–14 365, https://doi.org/10.1073/pnas.1318371111.

Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, https://doi.org/10.1175/BAMS-D-12-00050.1.

Larson, S. M., and B. P. Kirtman, 2014: The Pacific meridional mode as an ENSO precursor and predictor in the North American multimodel ensemble. *J. Climate*, **27**, 7018–7032, https://doi.org/10.1175/JCLI-D-14-00055.1.

Lin, H., and Coauthors, 2019: The Canadian Seasonal to Interannual Prediction System version 2 (CanSIPSv2). *Wea. Forcasting*, **35**, 1317–1343, https://doi.org/10.1175/WAF-D-19-0259.1.

Linkin, M. E., and S. Nigam, 2008: The North Pacific Oscillation–west Pacific teleconnection pattern: Mature-phase structure and winter impacts. *J. Climate*, **21**, 1979–1997, https://doi.org/10.1175/2007JCLI2048.1.

Liu, Z., and E. Di Lorenzo, 2018: Mechanisms and predictability of Pacific decadal variability. *Curr. Climate Change Rep.*, **4**, 128–144, https://doi.org/10.1007/s40641-018-0090-5.

Ma, J., S.-P. Xie, and H. Xu, 2017: Contributions of the North Pacific meridional mode to ensemble spread of ENSO prediction. *J. Climate*, **30**, 9167–9181, https://doi.org/10.1175/JCLI-D-17-0182.1.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, **78**, 1069–1080, https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2.

Merryfield, W. J., and Coauthors, 2013: The Canadian seasonal to interannual prediction system. Part I: Models and initialization. *Mon. Wea. Rev.*, **141**, 2910–2945, https://doi.org/10.1175/MWR-D-12-00216.1.

Miller, A. J., D. R. Cayan, T. P. Barnett, N. E. Graham, and J. M. Oberhuber, 1994: Interdecadal variability of the Pacific Ocean: Model response to observed heat flux and wind stress anomalies. *Climate Dyn.*, **9**, 287–302, https://doi.org/10.1007/BF00204744.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, https://doi.org/10.1175/2007JCLI1824.1.

Rogers, J. C., 1981: The North Pacific Oscillation. *J. Climatol.*, **1**, 39–57, https://doi.org/10.1002/joc.3370010106.

Saha, S., and Coauthors, 2014: The NCEP climate forecast system version 2. *J. Climate*, **27**, 2185–2208, https://doi.org/10.1175/JCLI-D-12-00823.1.

Sherman, K., and L. M. Alexander, Eds., 1986: *Variability and Management of Large Marine Ecosystems*. Avalon Publishing, 319 pp.

Stan, C., D. M. Straus, J. S. Frederiksen, H. Lin, E. D. Maloney, and C. Schumacher, 2017: Review of tropical-extratropical teleconnections on intraseasonal time scales. *Rev. Geophys.*, **55**, 902–937, https://doi.org/10.1002/2016RG000538.

Stock, C. A., and Coauthors, 2015: Seasonal sea surface temperature anomaly prediction for coastal ecosystems. *Prog. Oceanogr.*, **137**, 219–236, https://doi.org/10.1016/j.pocean.2015.06.007.

Stuecker, M. F., 2018: Revisiting the Pacific meridional mode. *Sci. Rep.*, **8**, 3216, https://doi.org/10.1038/s41598-018-21537-0.

Thomas, E. E., and D. J. Vimont, 2016: Modeling the mechanisms of linear and nonlinear ENSO responses to the Pacific meridional mode. *J. Climate*, **29**, 8745–8761, https://doi.org/10.1175/JCLI-D-16-0090.1.

Tommasi, D., and Coauthors, 2017: Managing living marine resources in a dynamic environment: The role of seasonal to decadal climate forecasts. *Prog. Oceanogr.*, **152**, 15–49, https://doi.org/10.1016/j.pocean.2016.12.011.

Trenberth, K. E., and J. W. Hurrell, 1994: Decadal atmosphere-ocean variations in the Pacific. *Climate Dyn.*, **9**, 303–319, https://doi.org/10.1007/BF00204745.

Vimont, D. J., D. S. Battisti, and A. C. Hirst, 2001: Footprinting: A seasonal connection between the tropics and mid-latitudes. *Geophys. Res. Lett.*, **28**, 3923–3926, https://doi.org/10.1029/2001GL013435.

——, J. M. Wallace, and D. S. Battisti, 2003: The seasonal footprinting mechanism in the Pacific: Implications for ENSO. *J. Climate*, **16**, 2668–2675, https://doi.org/10.1175/1520-0442(2003)016<2668:TSFMIT>2.0.CO;2.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109**, 784–812, https://doi.org/10.1175/1520-0493(1981)109<0784:TITGHF>2.0.CO;2.

Wen, C., Y. Xue, and A. Kumar, 2012: Seasonal prediction of North Pacific SSTs and PDO in the NCEP CFS hindcasts. *J. Climate*, **25**, 5689–5710, https://doi.org/10.1175/JCLI-D-11-00556.1.

Xie, S.-P., and S. G. H. Philander, 1994: A coupled ocean-atmosphere model of relevance to the ITCZ in the eastern Pacific. *Tellus*, **46A**, 340–350, https://doi.org/10.3402/tellusa.v46i4.15484.

Xu, T., M. Newman, A. Capotondi, and E. Di Lorenzo, 2021: The continuum of northeast Pacific marine heatwaves and their relationship to the tropical Pacific. *Geophys. Res. Lett.*, **48**, e2020GL090661, https://doi.org/10.1029/2020GL090661.

Yeager, S. G., and Coauthors, 2018: Predicting near-term changes in the Earth system: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model. *Bull. Amer. Meteor. Soc.*, **99**, 1867–1886, https://doi.org/10.1175/BAMS-D-17-0098.1.