Supporting Information for

Skillful Multiyear Sea Surface Temperature Predictability in CMIP6 Models and Historical Observations

Frances V. Davenport[1,2], Elizabeth A. Barnes[2], and Emily M. Gordon[2,3]

[1]Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, CO.
[2]Department of Atmospheric Science, Colorado State University, Fort Collins, CO.
[3]Department of Earth System Science, Stanford University, Stanford, CA.

Contents of this file

Hyperparameter tuning:

We tune the CNN hyperparameters using one GCM (*MPI-ESM1-2-LR*) and five locations across the globe (Fig S1a). The goal is to find a set of hyperparameters that performs well across all locations, and to then use the same architecture for all CNNs. We select hyperparameters sequentially using the following steps. In step 1, we tune the learning rate; in step 2, we tune the number of dense layers and neurons; in step 3, we tune the number of convolutional layers and filters; and in step 4, we tune the dropout rate and activity regularization parameter. At each step, we use keras tuner to train CNNs with different hyperparameter configurations. We then select a combination of hyperparameters that performs well on the validation set across all five locations before moving to the next step of tuning. In general, we find similarities in the best parameter combinations for each location, which supports our approach of using the same CNN architecture for all grid cells. However, it is possible that higher accuracy could be achieved in certain regions by tuning the architecture for that specific location, and therefore our results may slightly underestimate predictability. The results of the hyperparameter tuning are shown in Fig. S1. We found similar results when using different initial starting hyperparameters (results not shown).

CNN training:

We use a categorical cross-entropy loss function with the Adam optimizer, a batch size of 32, and define an epoch as 100 steps. The initial learning rate is 0.0003, and we use a learning rate scheduler to decrease the learning rate by a factor of $e^{-0.05}$ each epoch after the first 10 epochs. We use a dropout rate of 0.2 on the dense layer. We use early stopping to end training once the validation loss increases for at least 5 epochs. We train each CNN with three different random initializations, and we select the trained model that has lowest validation loss for later analyses.

.

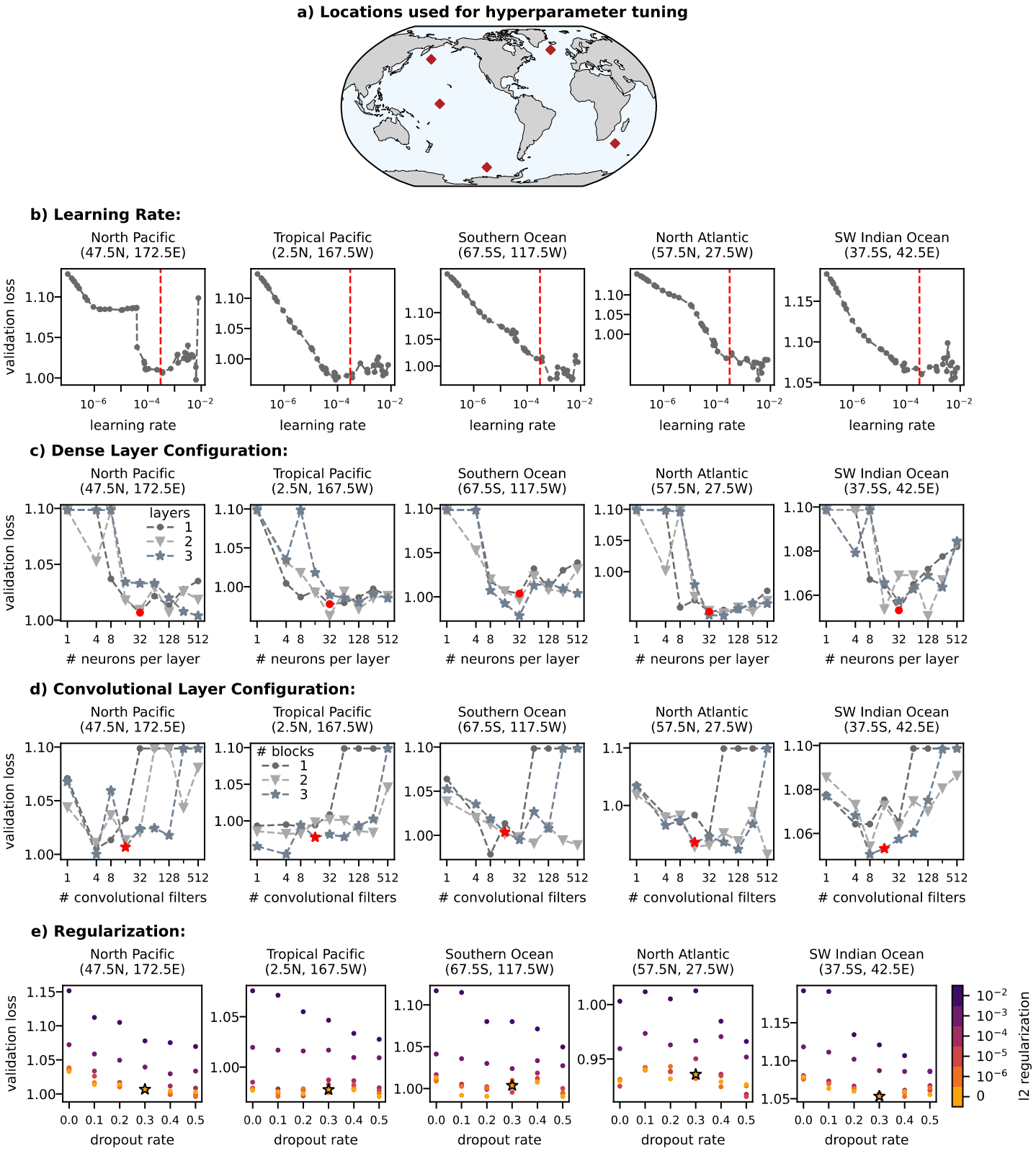

Figure S1. Hyperparameter tuning results. Selected parameters are shown by the red dashed line in b), red markers in c) and d), and the black stars in e).

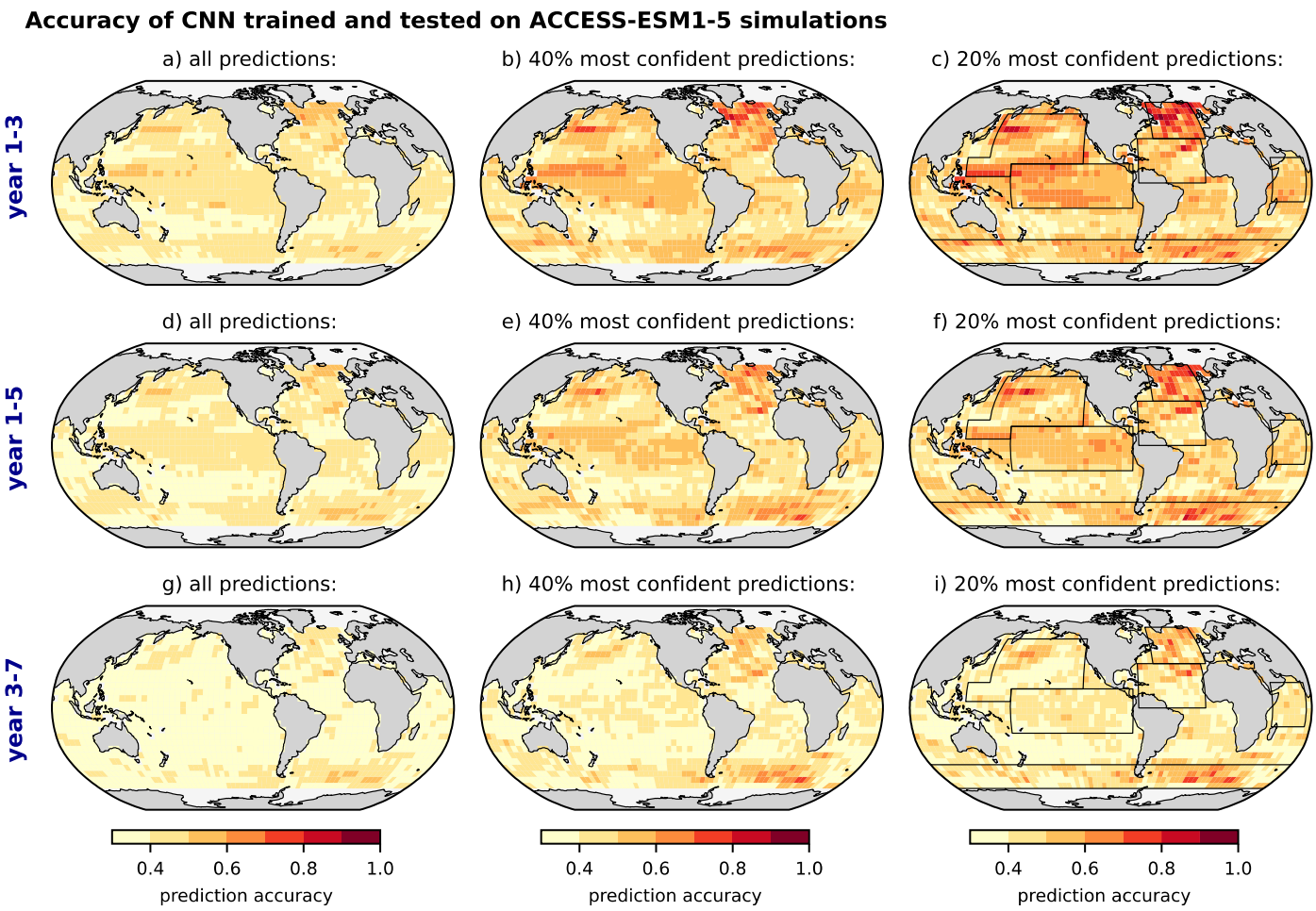

Figure S2. Same as Figure 2, but for ACCESS-ESM1-5.

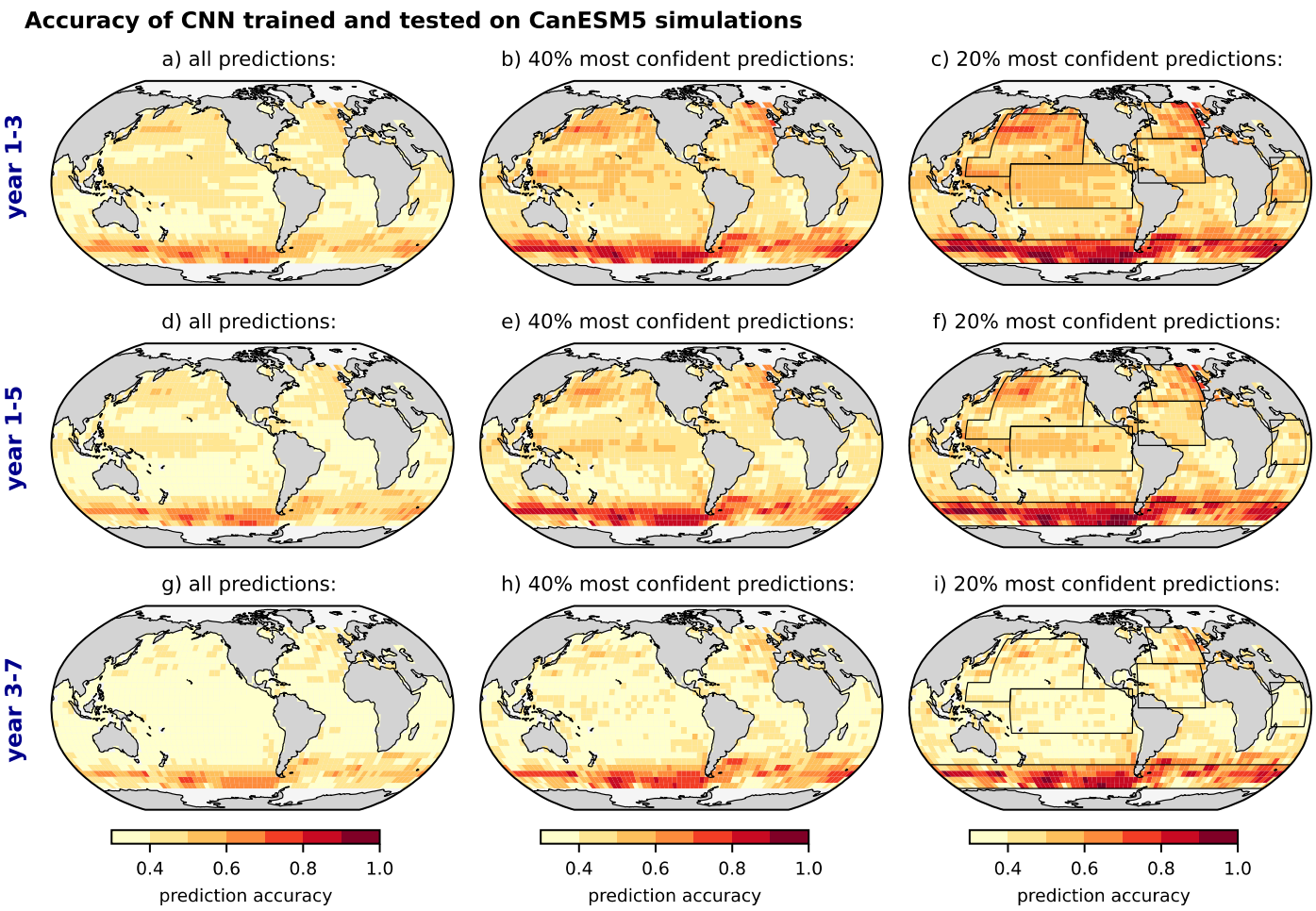

Figure S3. Same as Figure 2, but for CanESM5.

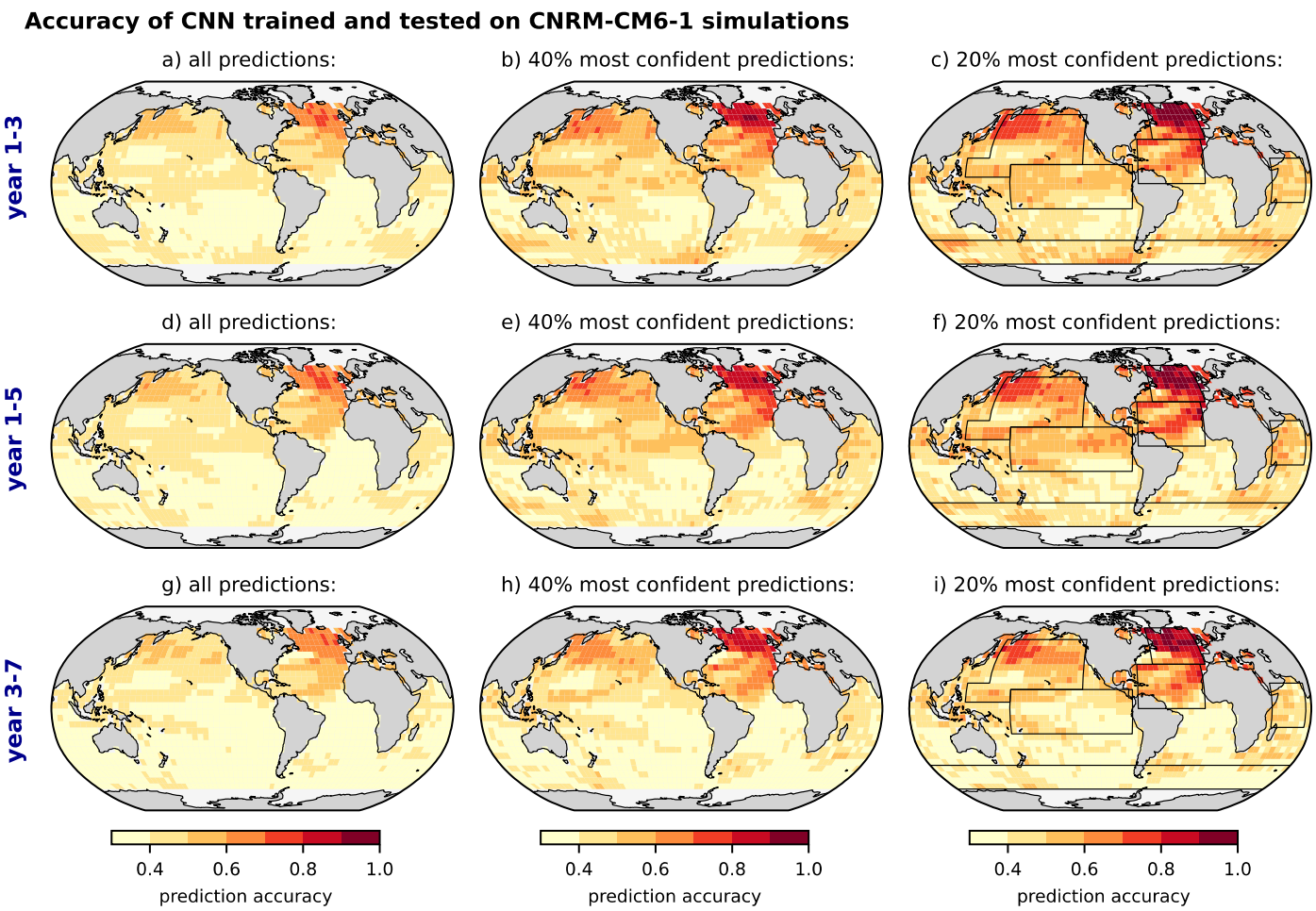

Figure S4. Same as Figure 2, but for CNRM-CM6-1.

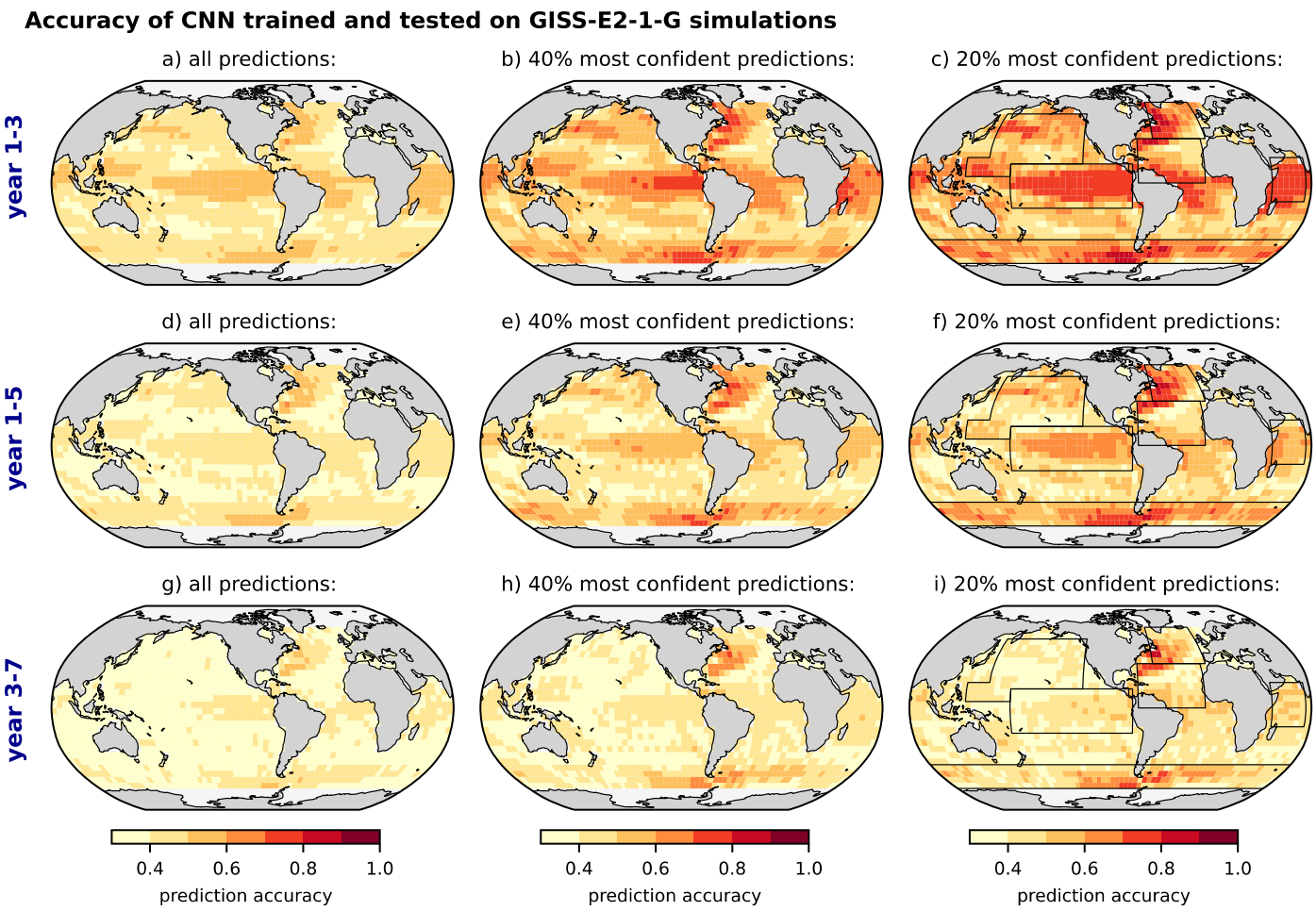

Figure S5. Same as Figure 2, but for GISS-E2-1-G.

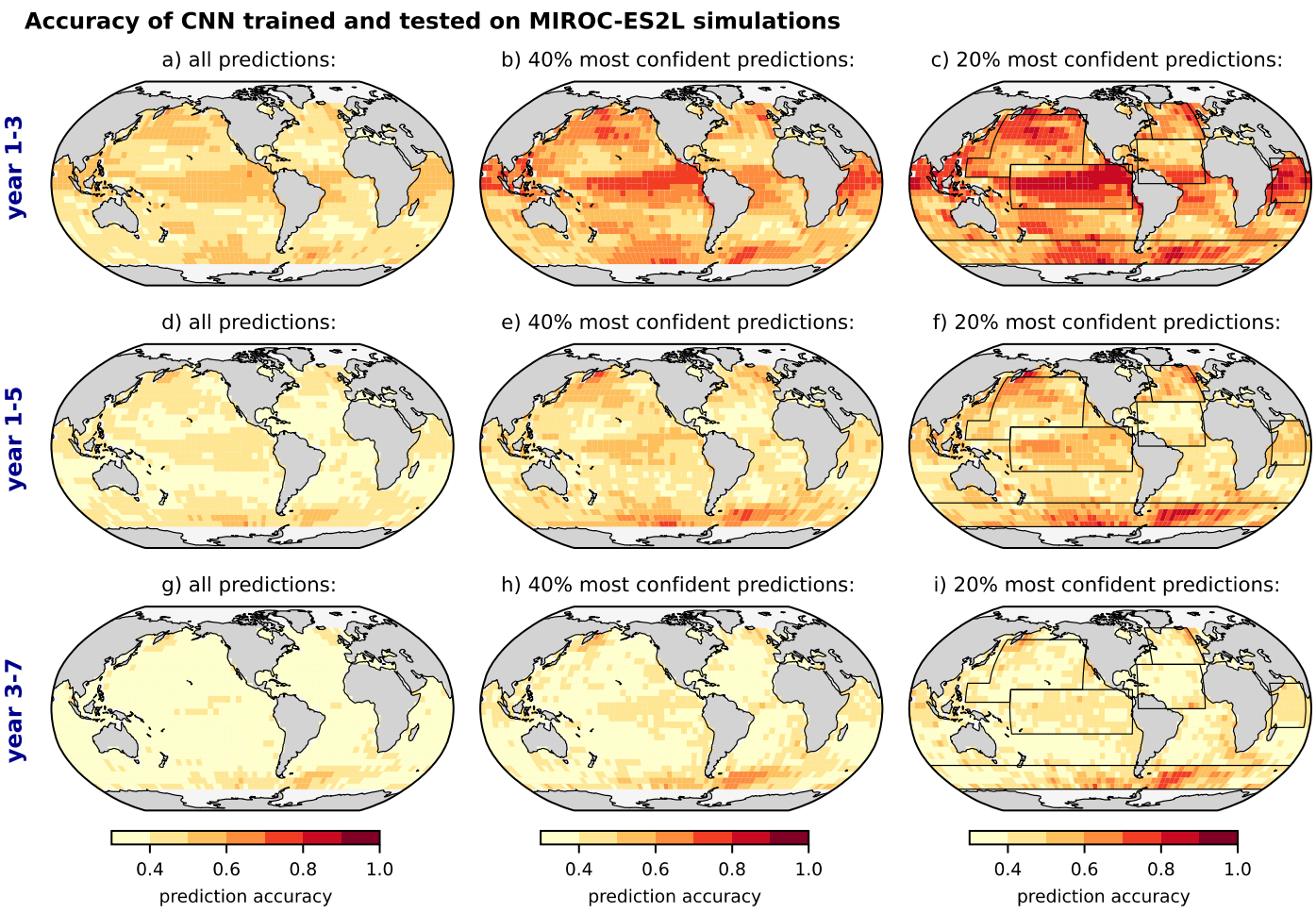

Figure S6. Same as Figure 2, but for MIROC-ES2L.

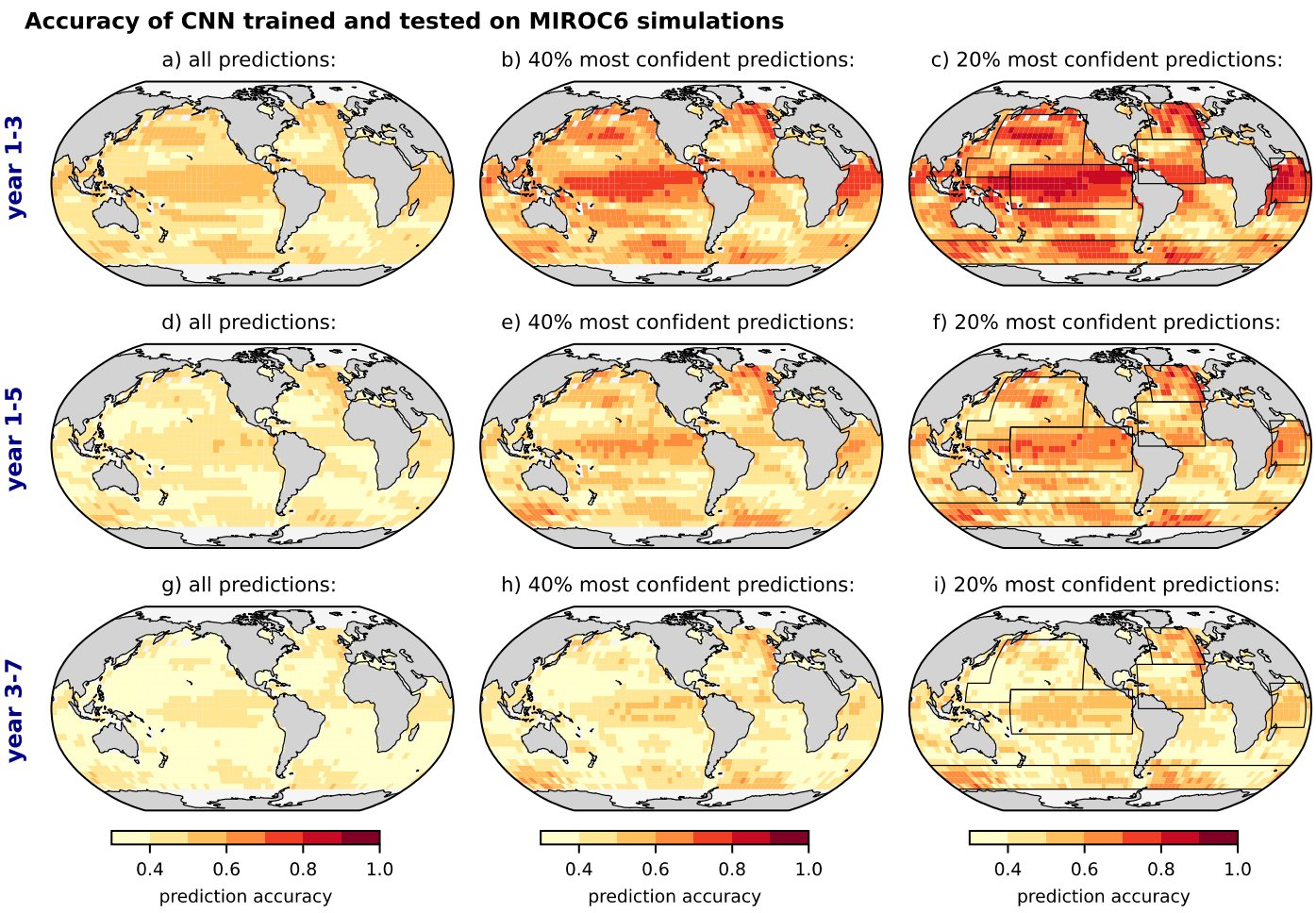

Figure S7. Same as Figure 2, but for MIROC6.

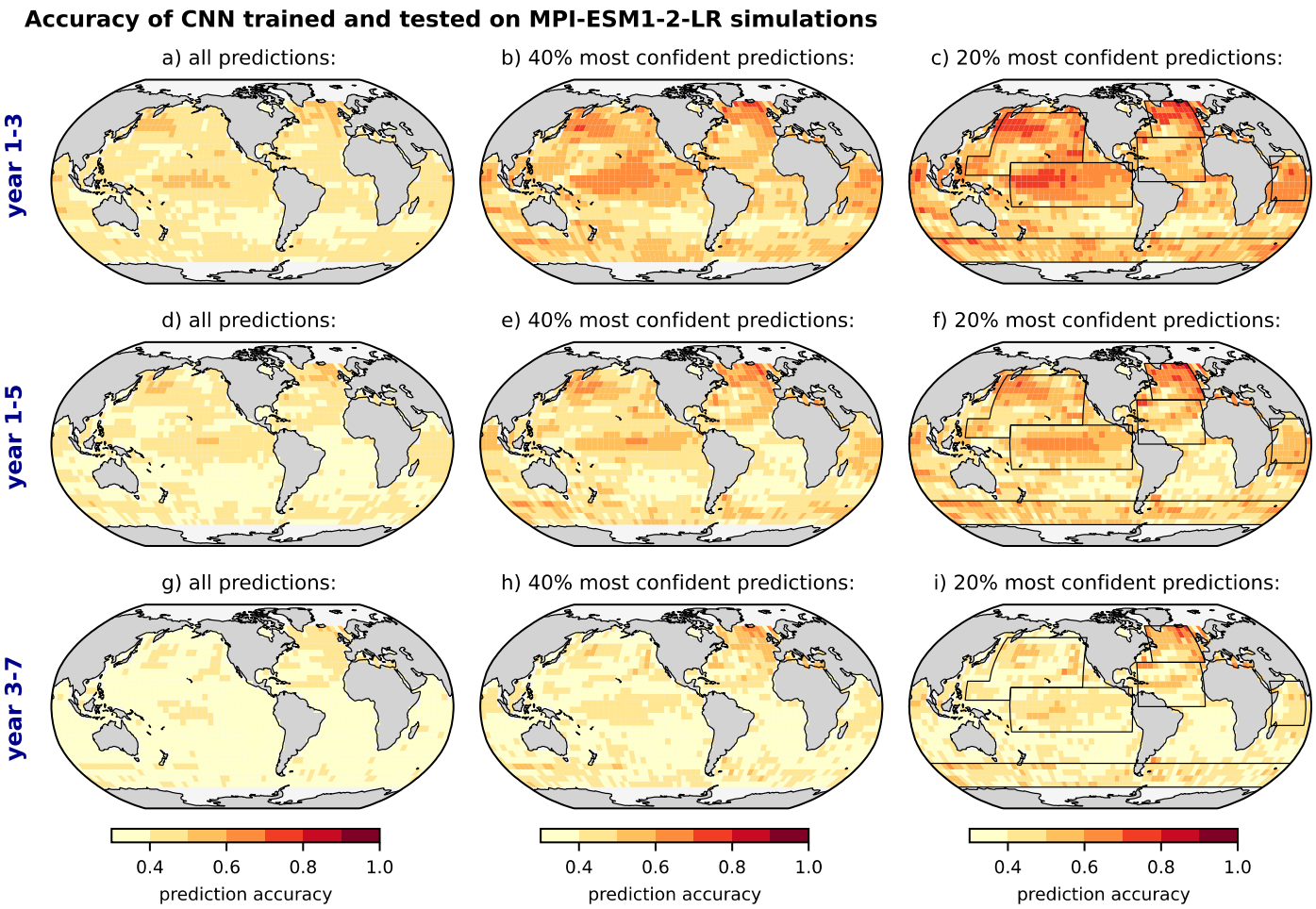

Figure S8. Same as Figure 2, but for MPI-ESM1-2-LR.

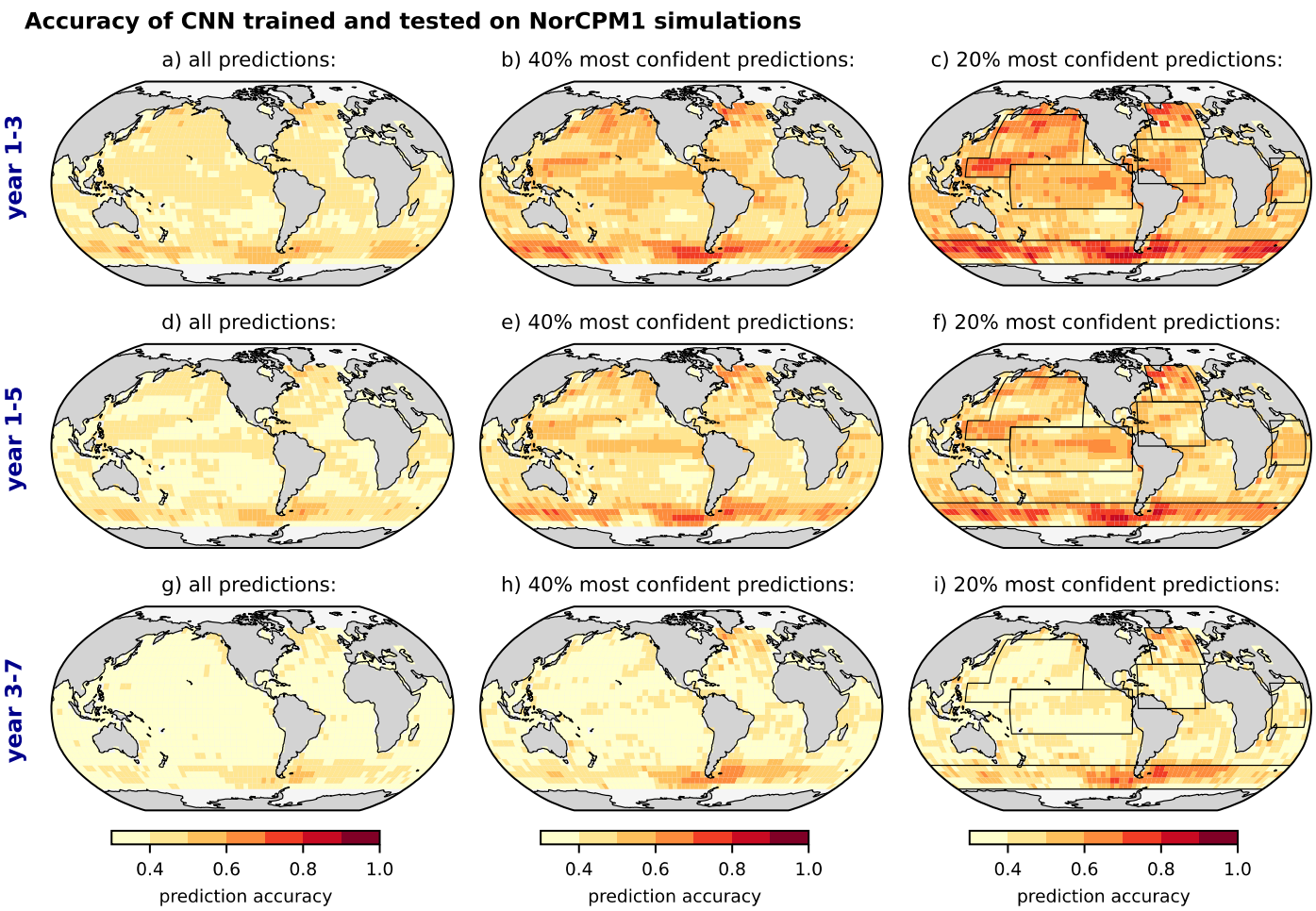

Figure S9. Same as Figure 2, but for NorCPM1.

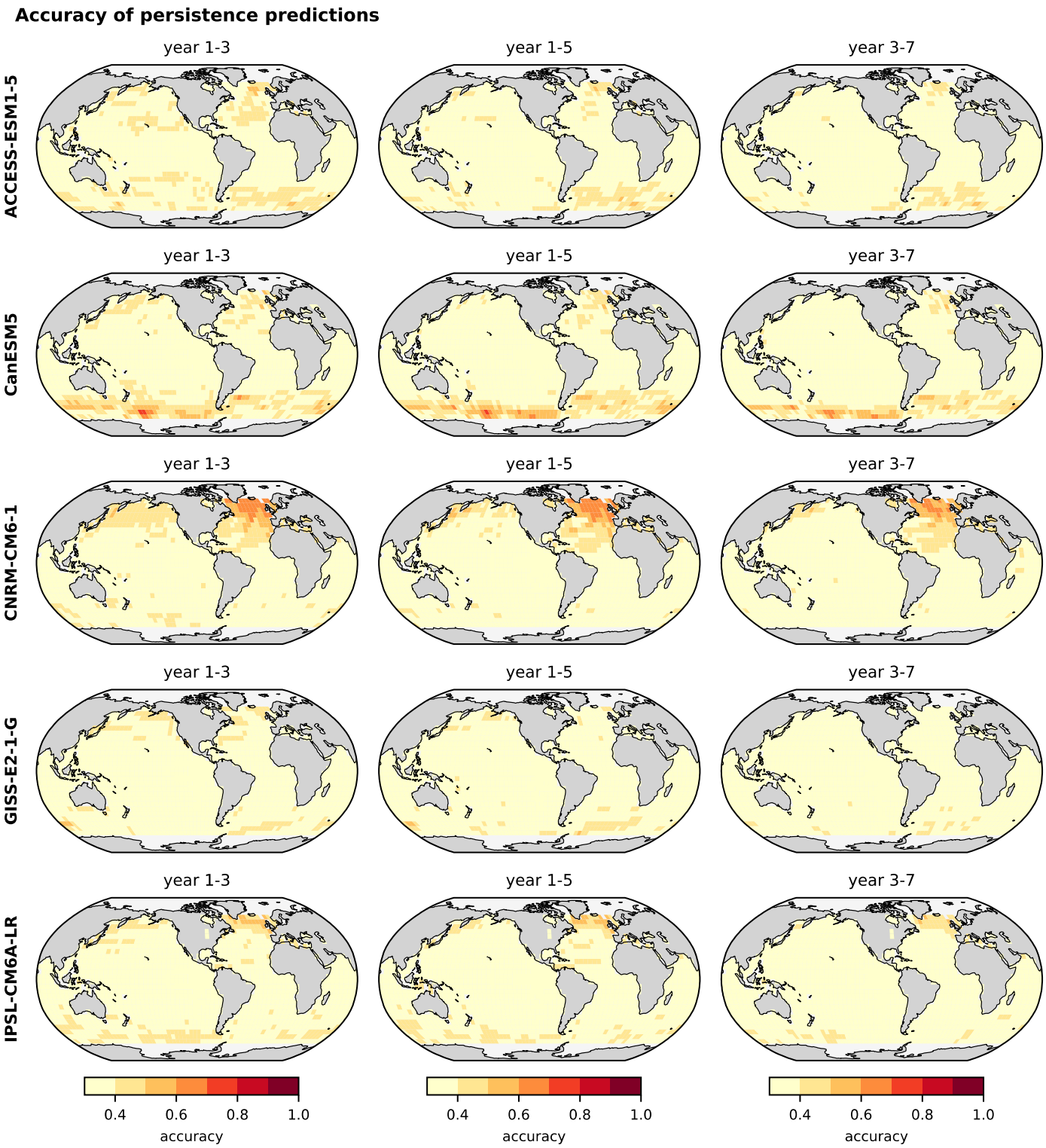

Figure S10. Accuracy of the persistence predictions for the three different lead times for ACCESS-ESM1-5, CanESM5, CNRM-CM6-1, GISS-E2-1-G, and IPSL-CM6A-LR.

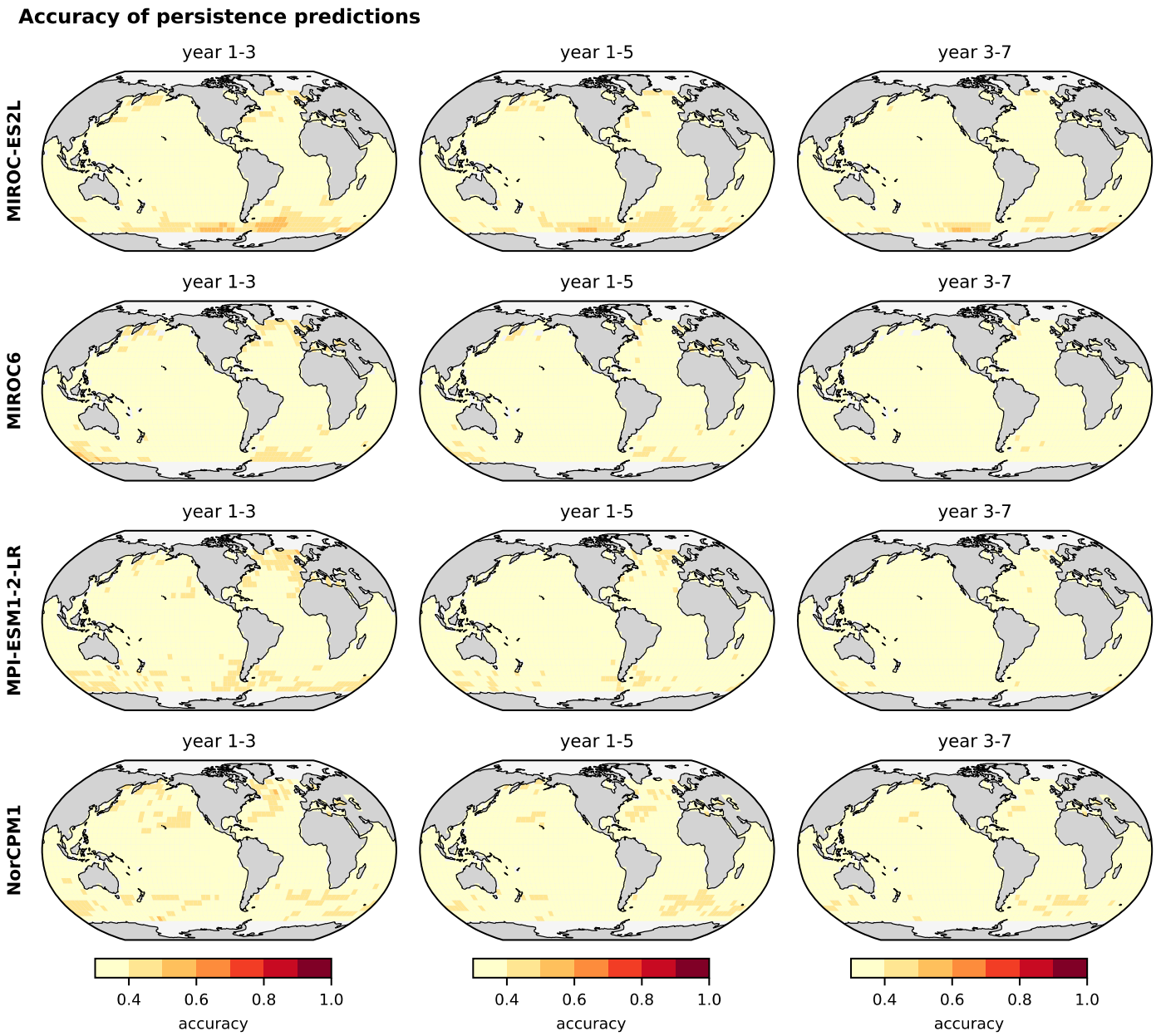

Figure S11. Same as Figure S10, but for MIROC-ES2L, MIROC6, MPI-ESM1-2-LR, NorCPM1.

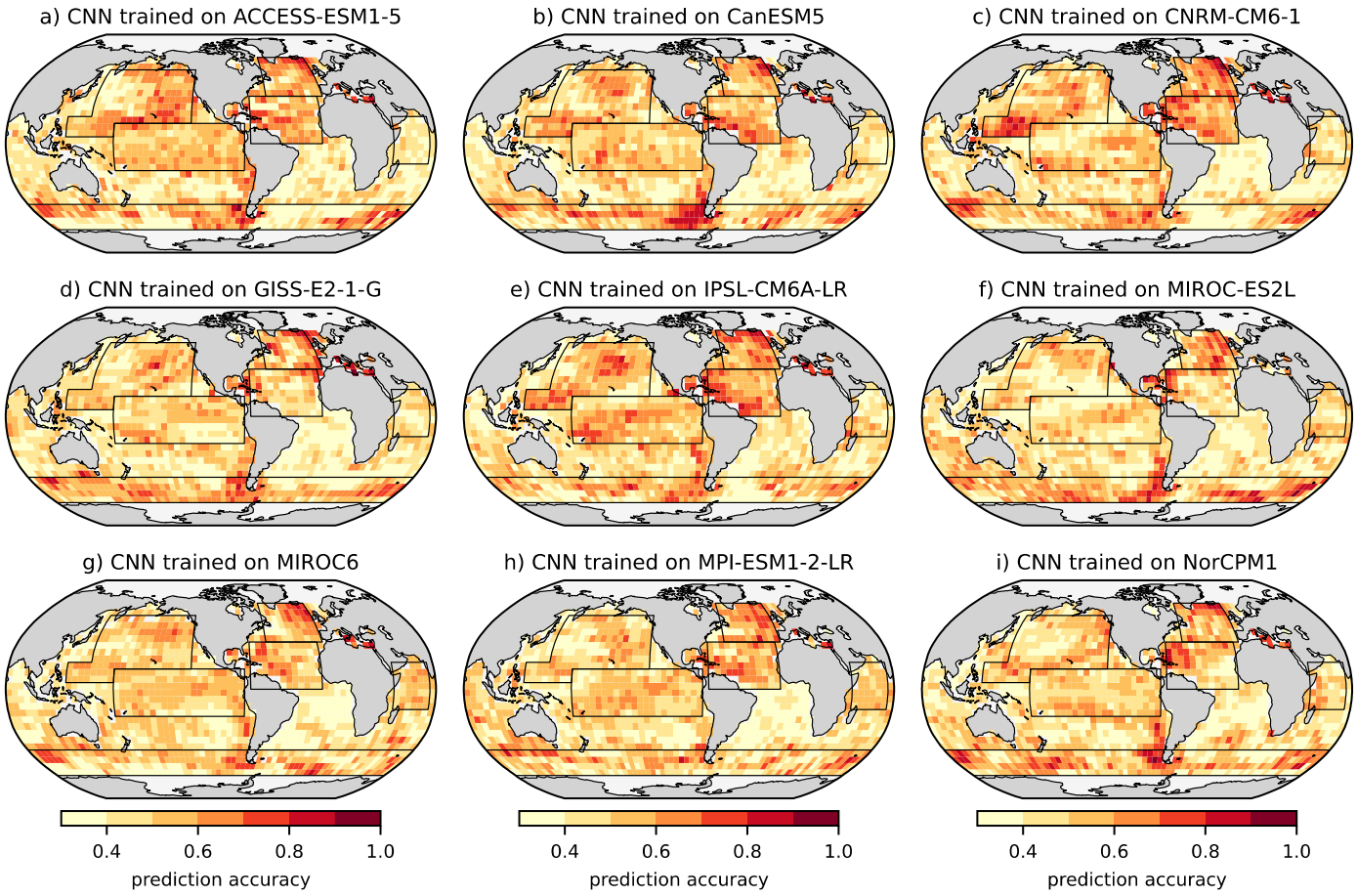

Figure S12. Same as Figure 3, but for year 1-3 predictions.

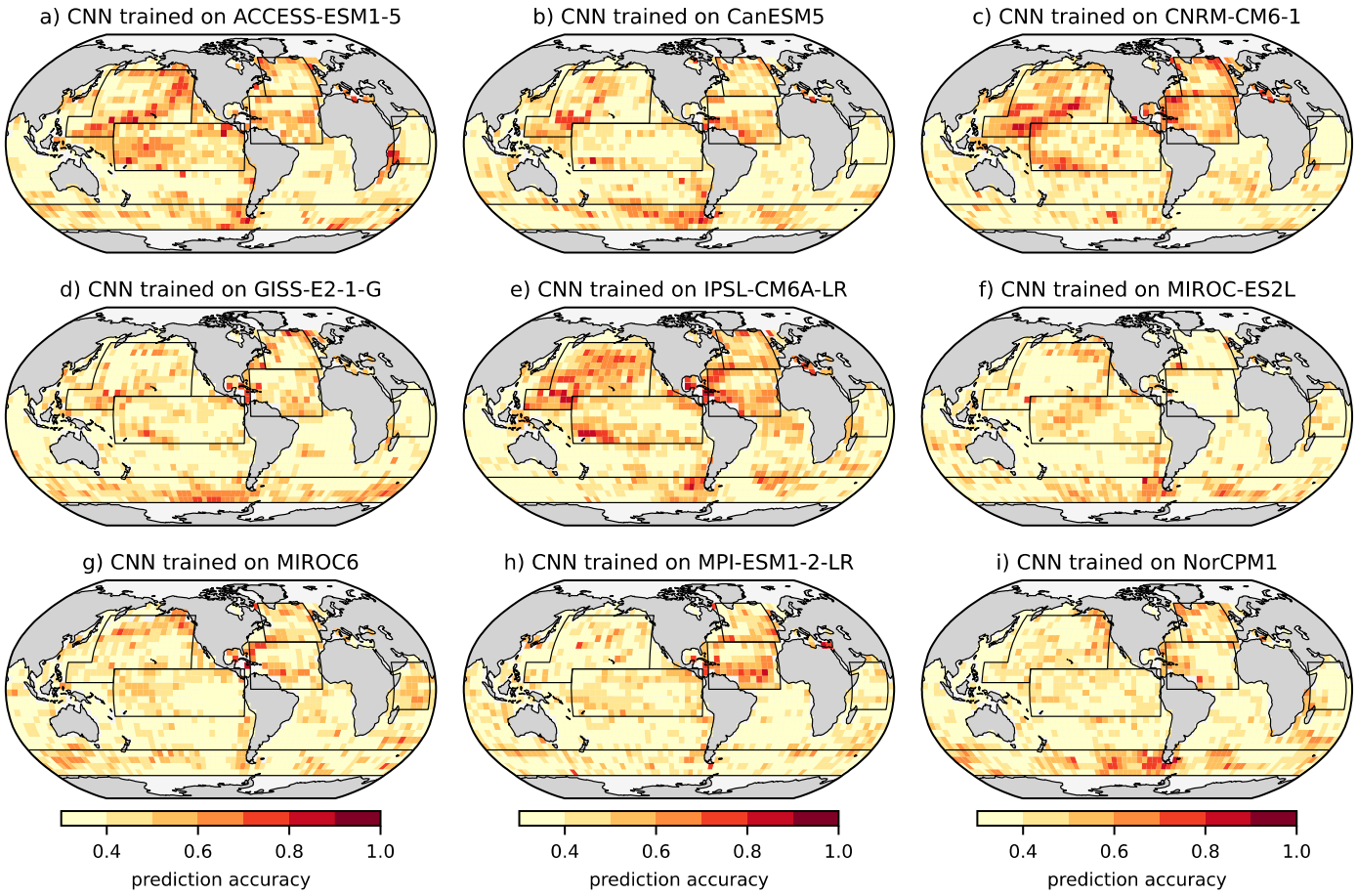

Figure S13. Same as Figure 3, but for year 3-7 predictions.

Accuracy of persistence predictions (ERSSTv5)

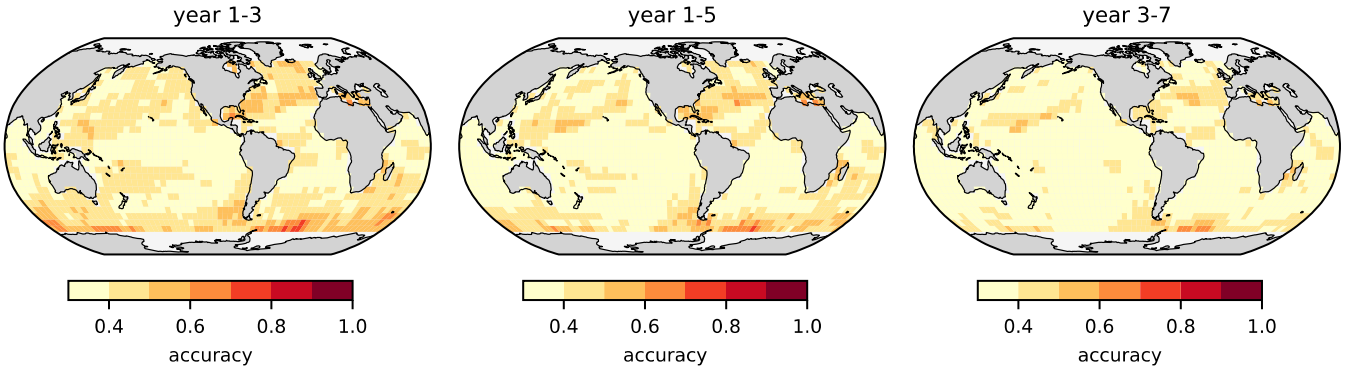

Figure S14. Accuracy of persistence predictions within ERSSTv5 observations for the three different lead times.

Table S1. Included CMIP6 models and simulations

Model	**Training** (22 simulations)	**Validation** (3 simulations)	**Testing** (5 simulations)
ACCESS-ESM1-5	r4i1p1f1, r5i1p1f1, r6i1p1f1, r7i1p1f1, r8i1p1f1, r9i1p1f1, r10i1p1f1, r11i1p1f1, r12i1p1f1, r13i1p1f1, r15i1p1f1, r16i1p1f1, r17i1p1f1, r19i1p1f1, r20i1p1f1, r22i1p1f1, r23i1p1f1, r24i1p1f1, r25i1p1f1, r26i1p1f1, r27i1p1f1, r30i1p1f1	r14i1p1f1, r21i1p1f1, r3i1p1f1	r18i1p1f1, r1i1p1f1, r28i1p1f1, r29i1p1f1, r2i1p1f1
CanESM5	r10i1p2f1, r11i1p2f1, r12i1p2f1, r13i1p2f1, r15i1p2f1, r16i1p2f1, r17i1p2f1, r19i1p2f1, r20i1p2f1, r22i1p2f1, r23i1p2f1, r24i1p2f1, r25i1p2f1, r26i1p2f1, r27i1p2f1, r30i1p2f1, r4i1p2f1, r5i1p2f1, r6i1p2f1, r7i1p2f1, r8i1p2f1, r9i1p2f1	r14i1p2f1, r21i1p2f1, r3i1p2f1	r18i1p2f1, r1i1p2f1, r28i1p2f1, r29i1p2f1, r2i1p2f1
CNRM-CM6-1	r10i1p1f2, r11i1p1f2, r12i1p1f2, r13i1p1f2, r15i1p1f2,r16i1p1f2, r17i1p1f2, r19i1p1f2, r20i1p1f2, r22i1p1f2, r23i1p1f2, r24i1p1f2, r25i1p1f2, r26i1p1f2, r27i1p1f2, r30i1p1f2, r4i1p1f2, r5i1p1f2, r6i1p1f2, r7i1p1f2, r8i1p1f2, r9i1p1f2	r14i1p1f2, r21i1p1f2, r3i1p1f2	r18i1p1f2, r1i1p1f2, r28i1p1f2, r29i1p1f2, r2i1p1f2
GISS-E2-1-G	r101i1p1f1, r102i1p1f1, r10i1p1f1, r10i1p3f1, r1i1p1f2,r1i1p3f1, r1i1p5f1, r2i1p1f2, r2i1p3f1, r3i1p3f1, r3i1p5f1, r4i1p1f1, r4i1p5f1, r5i1p1f1, r5i1p1f2, r5i1p3f1, r6i1p1f1, r6i1p3f1, r7i1p1f1, r8i1p1f1, r8i1p3f1, r9i1p1f1	r2i1p5f1, r3i1p1f1, r4i1p3f1	r1i1p1f1, r2i1p1f1, r3i1p1f2, r4i1p1f2, r9i1p3f1
IPSL-CM6A-LR	r10i1p1f1, r11i1p1f1, r12i1p1f1, r13i1p1f1, r15i1p1f1, r16i1p1f1, r17i1p1f1, r19i1p1f1, r20i1p1f1, r22i1p1f1, r23i1p1f1, r24i1p1f1, r25i1p1f1, r26i1p1f1, r27i1p1f1, r30i1p1f1, r4i1p1f1, r5i1p1f1, r6i1p1f1, r7i1p1f1, r8i1p1f1, r9i1p1f1	r14i1p1f1, r21i1p1f1, r3i1p1f1	r18i1p1f1, r1i1p1f1, r28i1p1f1, r29i1p1f1, r2i1p1f1
MIROC-ES2L	r10i1p1f2, r11i1p1f2, r12i1p1f2, r13i1p1f2, r15i1p1f2, r16i1p1f2, r17i1p1f2, r19i1p1f2, r20i1p1f2, r22i1p1f2, r23i1p1f2, r24i1p1f2, r25i1p1f2, r26i1p1f2, r27i1p1f2, r30i1p1f2, r4i1p1f2, r5i1p1f2, r6i1p1f2, r7i1p1f2, r8i1p1f2, r9i1p1f2	r14i1p1f2, r21i1p1f2, r3i1p1f2	r18i1p1f2, r1i1p1f2, r28i1p1f2, r29i1p1f2, r2i1p1f2
MIROC6	r10i1p1f1, r11i1p1f1, r12i1p1f1, r13i1p1f1, r15i1p1f1, r16i1p1f1, r17i1p1f1, r19i1p1f1, r20i1p1f1, r22i1p1f1, r23i1p1f1, r24i1p1f1, r25i1p1f1, r26i1p1f1, r27i1p1f1, r30i1p1f1, r4i1p1f1, r5i1p1f1, r6i1p1f1, r7i1p1f1, r8i1p1f1, r9i1p1f1	r14i1p1f1, r21i1p1f1, r3i1p1f1	r18i1p1f1, r1i1p1f1, r28i1p1f1, r29i1p1f1, r2i1p1f1
MPI-ESM1-2-LR	r10i1p1f1, r11i1p1f1, r12i1p1f1, r13i1p1f1, r15i1p1f1,r16i1p1f1, r17i1p1f1, r19i1p1f1, r20i1p1f1, r22i1p1f1, r23i1p1f1, r24i1p1f1, r25i1p1f1, r26i1p1f1, r27i1p1f1, r30i1p1f1, r4i1p1f1, r5i1p1f1, r6i1p1f1, r7i1p1f1, r8i1p1f1, r9i1p1f1	r14i1p1f1, r21i1p1f1, r3i1p1f1	r18i1p1f1, r1i1p1f1, r28i1p1f1, r29i1p1f1, r2i1p1f1
NorCPM1	r10i1p1f1, r11i1p1f1, r12i1p1f1, r13i1p1f1, r15i1p1f1, r16i1p1f1, r17i1p1f1, r19i1p1f1, r20i1p1f1, r22i1p1f1, r23i1p1f1, r24i1p1f1, r25i1p1f1, r26i1p1f1, r27i1p1f1, r30i1p1f1, r4i1p1f1, r5i1p1f1, r6i1p1f1, r7i1p1f1, r8i1p1f1, r9i1p1f1	r14i1p1f1, r21i1p1f1, r3i1p1f1	r18i1p1f1, r1i1p1f1, r28i1p1f1, r29i1p1f1, r2i1p1f1