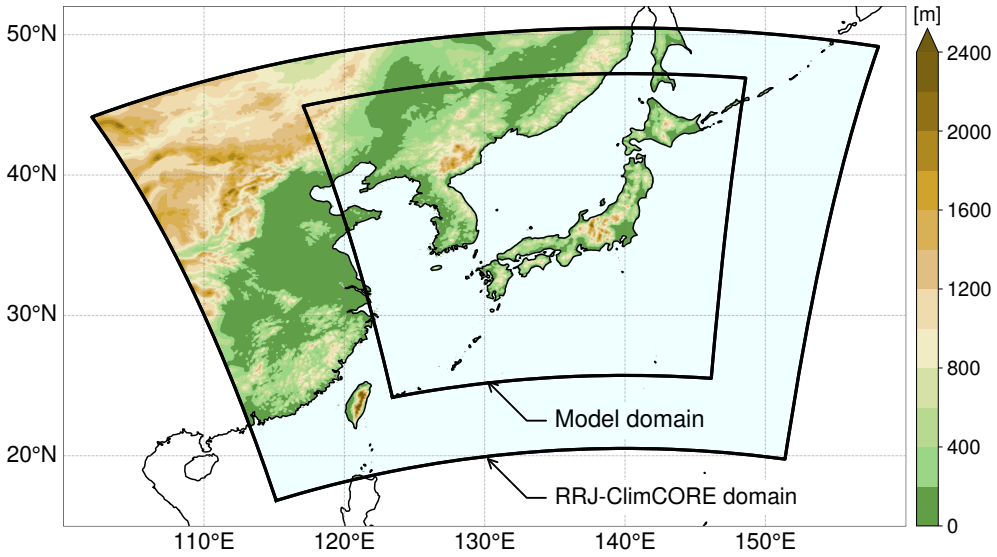

Fig. S1 Domains of the original RRJ-ClimCORE dataset and the subdomain extracted for input and output of our MLWP model. Colors indicate the surface elevation from RRJ-ClimCORE.

Table S1 Parameter settings of the loss functions for each model. In the $\mathcal{V}_*$ rows, each cell lists the prognostic variables in the corresponding $\mathcal{V}_*$; $\emptyset$ indicates that the set contains no variables. The hyperparameters η_* and ξ are defined based on the use of units of m s^{-1} for wind speed and km^{-1} for gradients. The coefficients λ_* were adopted after confirming, during training of the *all* model, that each loss component attained comparable orders of magnitude on the validation data, thereby avoiding dominance by any single term.

Parameters	**mse**	**wind**	**wind+falfcl**	**wind+grad**	**all**
$\mathcal{V}_{\mathrm{mse}}$	SAT, SLP, TPW, DLR, DSR, T_p,Q_p,Z_p, U_p,V_p,$\mathrm{U}_{\mathrm{sfc}}$,$\mathrm{V}_{\mathrm{sfc}}$, R1h	SAT, SLP, TPW, DLR, DSR, T_p,Q_p,Z_p, R1h	SAT, SLP, TPW, DLR, DSR, T_p,Q_p,Z_p	SAT, SLP, TPW, DLR, DSR, T_p,Q_p,Z_p, R1h	SAT, SLP, TPW, DLR, DSR, T_p,Q_p,Z_p
$\mathcal{V}_{\mathrm{falfcl}}$	$\emptyset$	$\emptyset$	R1h	$\emptyset$	R1h
$\mathcal{V}_{\mathrm{grad}}$	$\emptyset$	$\emptyset$	$\emptyset$	SLP,Z_p	SLP,Z_p
λ_{mse}	1.0	1.0	1.0	1.0	1.0
λ_{wind}	0	1.0	1.0	1.0	1.0
$\lambda_{\mathrm{falfcl}}$	0	0	1.0	0	1.0
λ_{grad}	0	0	0	1.0	1.0
$w_{\mathrm{sfc}}, w_{1000}$	-	1.0, 1.0	1.0, 1.0	1.0, 1.0	1.0, 1.0
$w_{850}, w_{500}, w_{300}$	-	0.5, 0.3, 0.2	0.5, 0.3, 0.2	0.5, 0.3, 0.2	0.5, 0.3, 0.2
$\eta_{\mathrm{mag}}, \eta_{\mathrm{div}}, \eta_{\mathrm{curl}}$	-	1.0, 1.0, 1.0	1.0, 1.0, 1.0	1.0, 1.0, 1.0	1.0, 1.0, 1.0
ζ	-	-	1.0	-	1.0
ξ	-	-	-	500.0	500.0

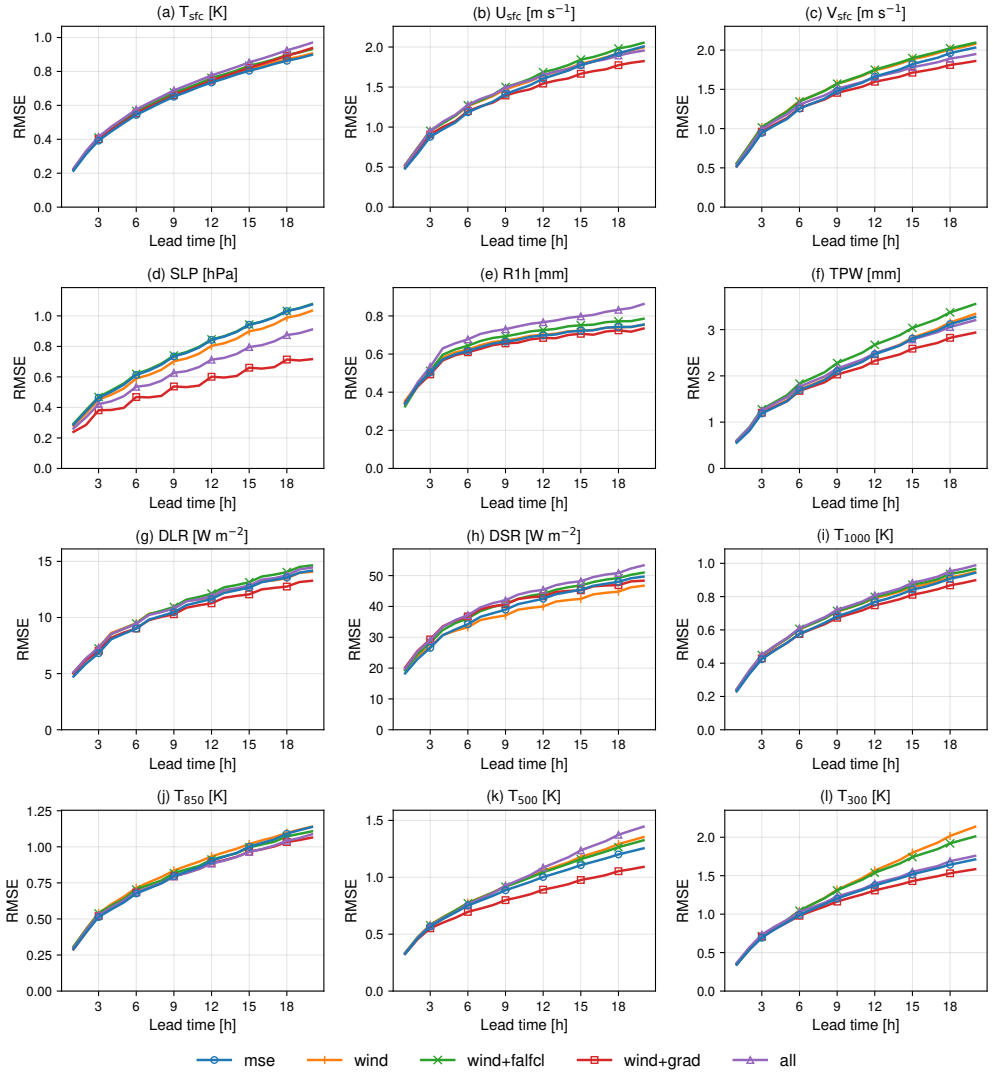

Fig. S2 RMSE of variables from five models (blue: mse, orange: wind, green: wind+falfcl, red: wind+grad, purple: all) over a 12-month period in 2022, shown as a function of forecast lead time. The variables are (a) T_{sfc}, (b) U_{sfc}, (c) V_{sfc}, (d) SLP, (e) R1h, (f) TPW, (g) DLR, (h) DSR, (i) T_{1000}, (j) T_{850}, (k) T_{500}, and (l) T_{300}. Panels (b) and (d) are identical to Figs. 1d and 1a, respectively, and are included here for ease of reference.

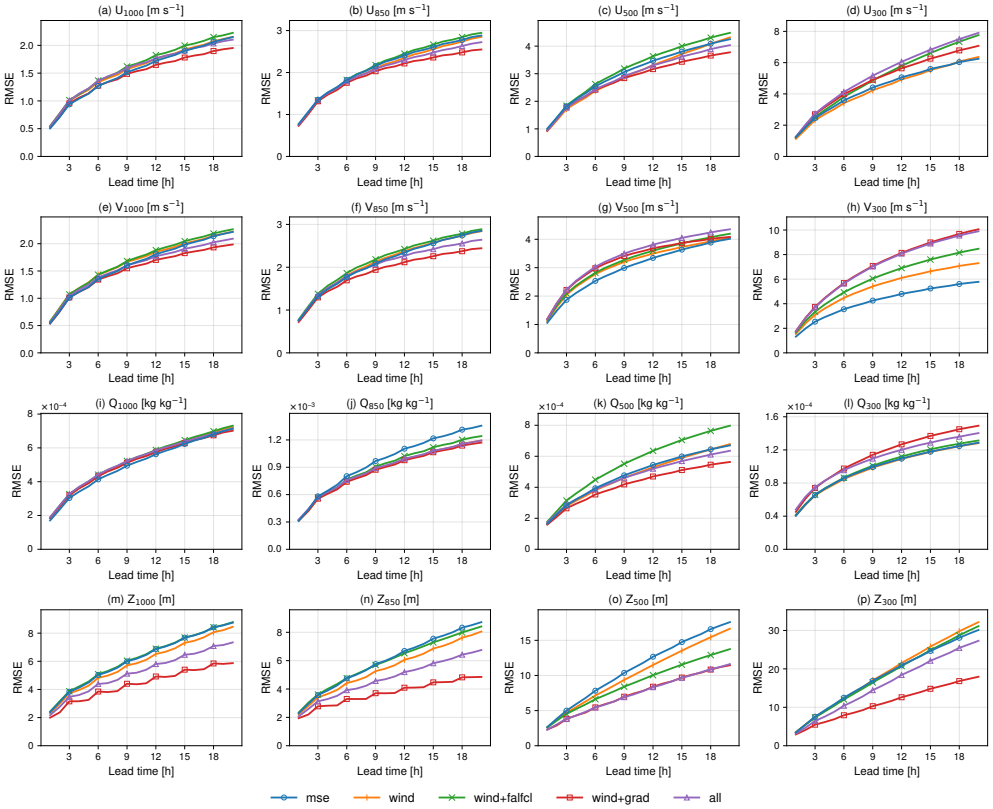

Fig. S3 Same as Fig. S2, but for (a) U_{1000}, (b) U_{850}, (c) U_{500}, (d) U_{300}, (e) V_{1000}, (f) V_{850}, (g) V_{500}, (h) V_{300}, (i) Q_{1000}, (j) Q_{850}, (k) Q_{500}, (l) Q_{300}, (m) Z_{1000}, (n) Z_{850}, (o) Z_{500}, and (p) Z_{300}. Panels (c), (d), (n) and (o) are identical to Figs. 1d, 1e, 1b, and 1c respectively, and are included here for ease of reference.

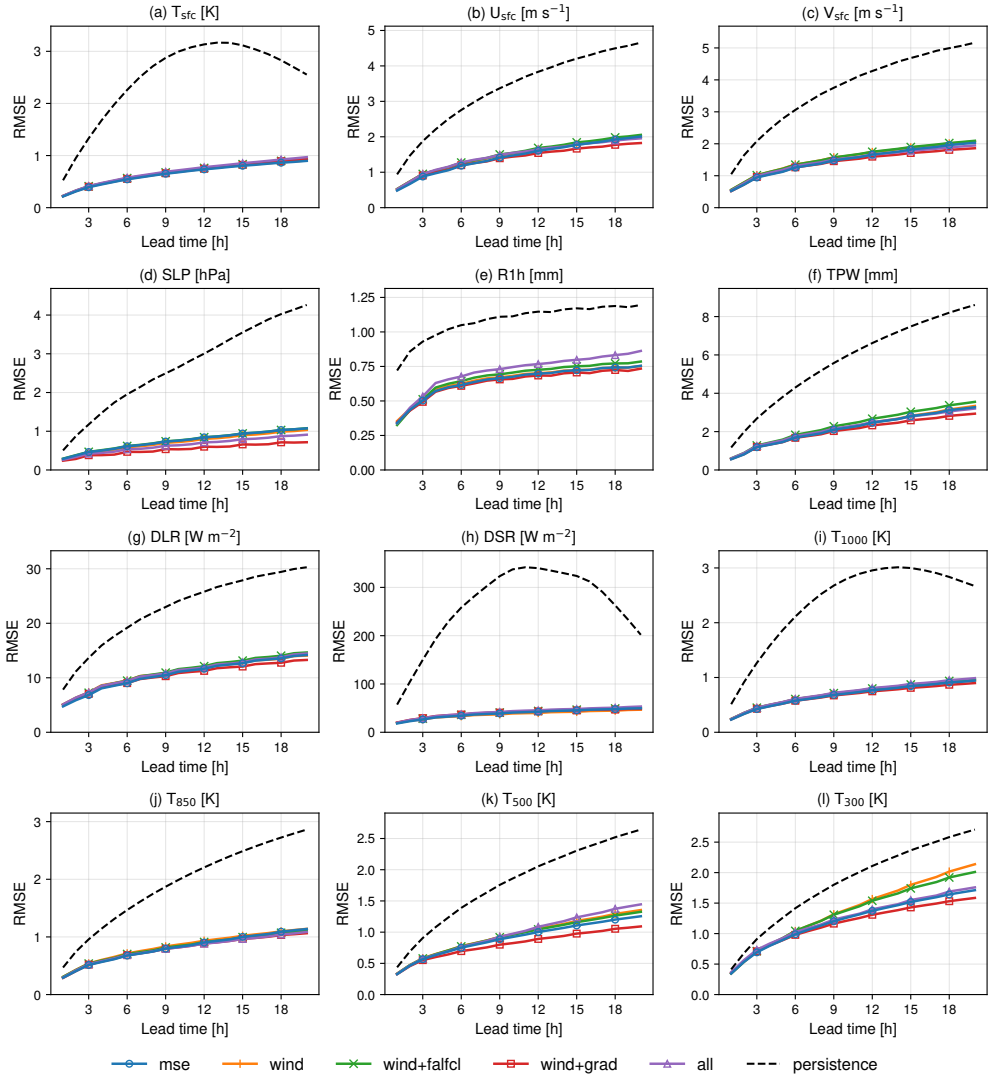

Fig. S4 Same as Fig. S2, but including a persistence baseline (black dashed line), which assumes the current state persists so that future weather equals the initial state.

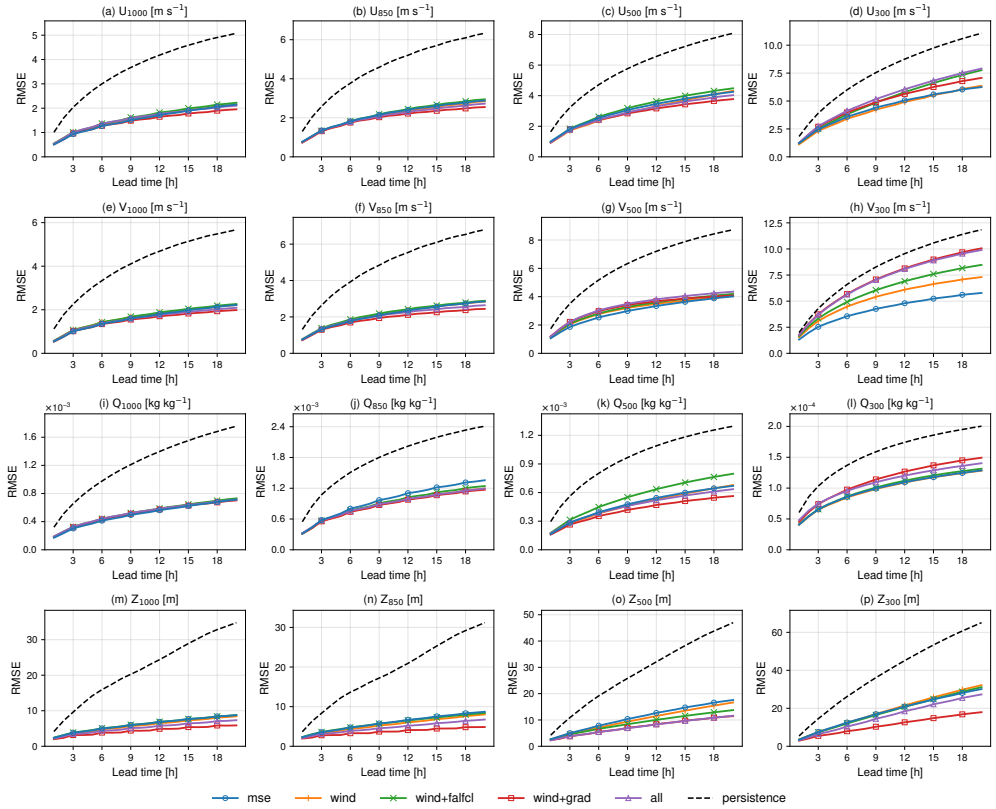

Fig. S5 Same as Fig. S3, but including a persistence baseline (black dashed line), which assumes the current state persists so that future weather equals the initial state.

Table S2 Central SLP and maximum surface wind speed of Typhoon Nanmadol from the best track, reanalyses (RRJ-ClimCORE and ERA5), and 6-hour forecasts of each model (mse, wind, wind+falfcl, wind+grad, all, and AIFS). Forecasts are initialized at 00, 03, and 06 UTC on 18 September 2022, with corresponding valid times at 06, 09, and 12 UTC, respectively.

			Reanalyses		MLWP models (FT=6h)					
	valid (UTC)	Best track	RRJ-ClimCORE	ERA5	**mse**	**wind**	**wind +falfcl**	**wind +grad**	**all**	AIFS
SLP [hPa]	06	930.0	934.6	955.3	943.4	936.4	951.5	943.3	948.9	952.6
	09	940.0	939.2	958.6	948.1	939.9	951.9	952.1	953.9	–
	12	955.0	949.4	962.9	952.3	945.4	957.5	954.3	955.0	962.8
Wind speed [m s^{-1}]	06	46.3	37.7	28.3	34.1	38.8	30.3	32.5	30.4	26.1
	09	43.7	32.1	27.7	31.8	33.5	28.7	27.4	26.9	–
	12	38.6	29.8	25.9	30.7	32.9	26.7	28.5	28.5	24.9

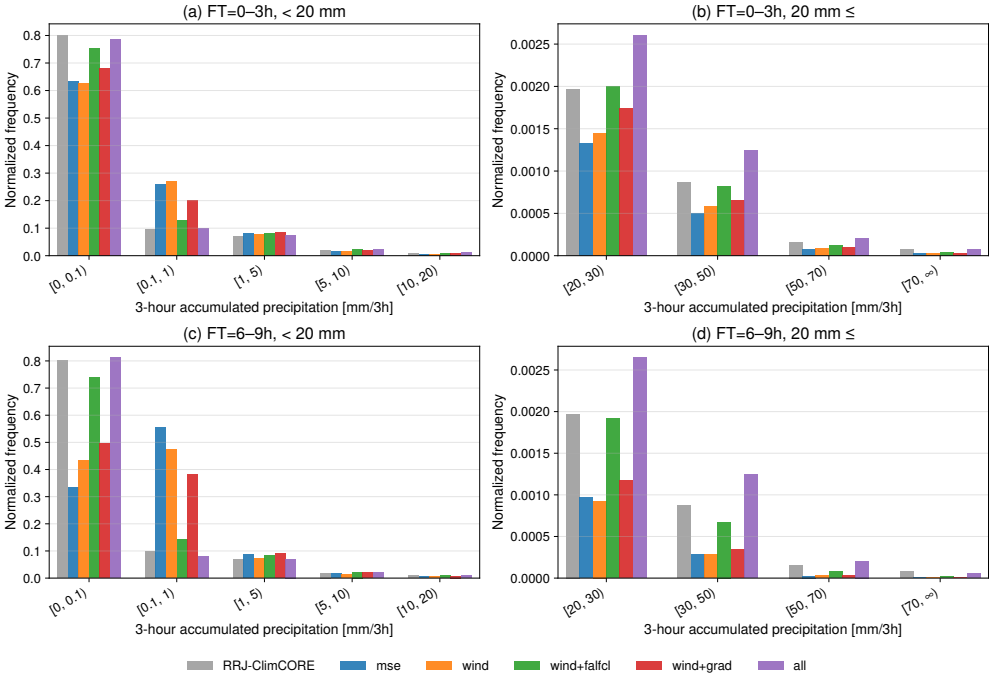

Fig. S6 Normalized frequency distributions of 3-hour accumulated precipitation from five models (blue: mse, orange: wind, green: wind+falfcl, red: wind+grad, purple: all) and RRJ-ClimCORE (gray, ground truth) for two forecast ranges in 2022. Panels (a)–(b) show forecasts for FT=0–3 h and (c)–(d) for FT=6–9 h. The precipitation bins are divided into (a,c) below 20 mm/3h and (b,d) above 20 mm/3h. Frequencies are computed by counting precipitation amounts at each grid point over the domain.