Figure S1. Global distributions of 30-h Ellrod1 forecasts at 250 hPa calculated by using seven TIGGE NWP model outputs for a 6-month period (2016.10–2017.03).
Figure S2. The same as Fig. S1, except for 30-h DVT forecasts.
Figure S3. PODY-POFD lines with the AUC skill scores (lower right) for 30-h deterministic (a) 1/Ri and (b) NCSU1 forecasts derived from seven TIGGE NWP models. The single-model-based forecast and MMEM forecast are depicted by dashed and thick solid lines, respectively. The evaluation period is six months (2016.10–2017.03). For the evaluation, NIL and MOG turbulence observations reported from in situ EDR and PIREPs are used. The diagonal line is the no-skill line.
Figure S4. Reliability diagrams (thick solid lines) for three 30-h MMP forecasts based on (a) $1/Ri$ and (b) NCSU1 diagnostics derived from seven TIGGE NWP model outputs. The evaluation period is six months (2016.10–2017.03). The three MMP forecasts are derived by applying the threshold of the 95th- (black line), 98th- (red line), and 99th- (blue line) percentile values of seven PDFs of $1/Ri$ and NCSU1. The PDF is constructed using values of $1/Ri$ and NCSU1 at 250 hPa in the 60°S–60°N latitude band for the evaluation period. The numbers of NIL (gray bars) and MOG (white tips on top of gray bars) turbulence events captured in each forecast probability of the 95th- (dark gray), 98th- (gray), and 99th- (light gray) percentile value-based MMP forecasts are depicted by bar chart. For better comparison, the numbers of NIL and MOG turbulence events in the 0% probability are multiplied by 1/5. The diagonal (black thin) and the horizontal (black dotted) lines represent the perfect reliability and the sample climatology lines, respectively.