weathertrends360

better business in any weather®

A Comparison of Forecasting Skill Between Weather Trends International (WTI) and The Climate Prediction Center (CPC) at The National Oceanic and Atmospheric Administration (NOAA)

Laurence S. Kalkstein, Ph.D.
Climatologist and Professor Emeritus, President
Applied Climatologists, Inc.
Chief Heat Science Advisor
Arsht-Rockefeller Foundation

INTRODUCTION

The purpose of this analysis is to compare the accuracy of three-month temperature and precipitation forecasts that are issued by the Climate Prediction Center (CPC) against the accuracy of similar forecasts issued by Weather Trends International (WTI). The comparison extends for 182 three month forecast periods, from November/December/January, 2006-07 through December/January/February, 2021-2022.

These forecasts were developed for 260 cities around the country, and are issued by WTI at the end of the calendar month for the time period that is valid 11 months into the future. For example, the May, 2010 forecast was issued by WTI in late June, 2009. They are issued one month at a time, and three months are combined for the purposes of this analysis and to match up to CPC's forecast period. Once the forecasts are issued, they are not updated by WTI.

The CPC's forecasts are constructed differently; they are developed for three combined month intervals, but in this analysis, the CPC forecast developed one month prior to the three-month period was used in the comparison. Thus, for example, the May, 2010 CPC forecast was utilized in this comparison to evaluate their June, July, and August, 2010 forecast. Unlike the WTI forecasts, this permits the CPC to alter their forecasts through the time period, right up to the month before. All things being equal, this would offer an advantage to the accuracy of the CPC forecast as compared to WTI's.

During this evaluation period, 30-year normals were changed by NOAA every 10 years, and this is reflected in the data with normal changes occurring at the beginning of 2011 and 2021.

"Skill scores" were calculated by WTI for both sets of forecasts for each three-month period. The CPC has historically utilized the Heidke Skill Score (http://www.cpc.ncep.noaa.gov/products/Soilmst_Monitoring/US/Outlook/cas_score.shtml), which is defined as follows:

HSS (%) = 100 * (H - E) / (T - E)

where H = number of correct forecasts, E = expected number of correct forecasts (1/3 of total), and T = total number of valid forecast-observation pairs. An HSS above 0 translates to more "hits" than "misses" among the 260 forecasts for each three month period; conversely, an HSS below 0 indicates more misses than hits. An HSS of 0 would be, in the long term, the random result if there were no forecasting skill.

There are three categories utilized by CPC in forecasting: above average, below average, and average (the CPC defines these as "above normal", "normal", and "below normal" (CPC, 2016; http://www.cpc.ncep.noaa.gov/products/predictions/long_range/tools.html)). Since "normal", as

defined by the CPC, means an equal chance (EC) of above normal, normal, or below normal, we have instead utilized a 2-class system of above and below average, thus eliminating the EC, which is basically no forecast at all. Therefore, the area in white within Figure 1 is not included in the comparison, since there is a 33%-33%-33% chance of each of the three categories in that area. Thus, both the WTI and CPC forecast evaluation are based on a two category system, above and below average, to keep the forecasts parallel.
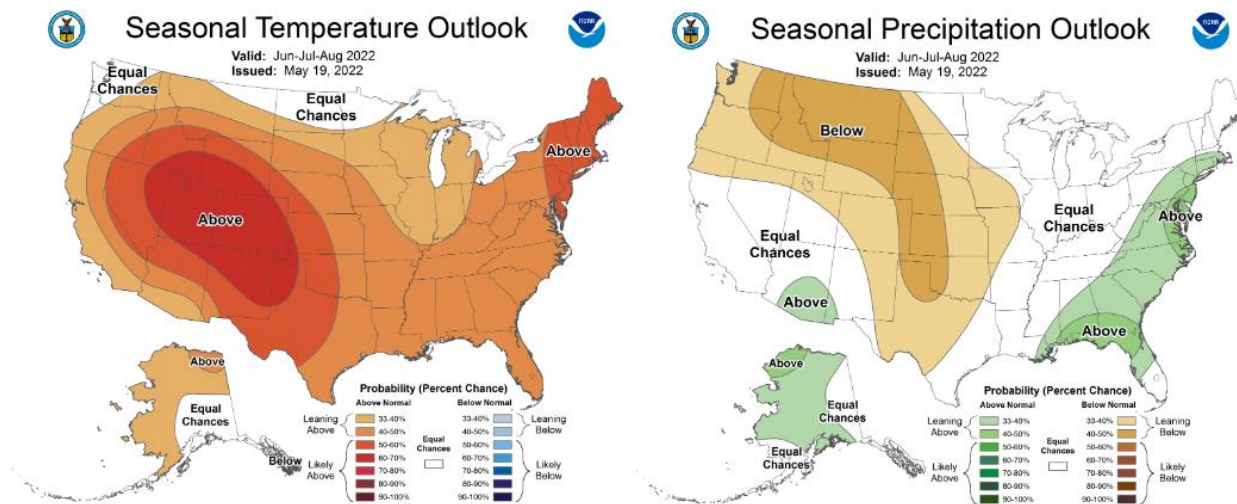
Figure 1. CPC seasonal forecast for Jun-Aug 2022, constructed 19 May 2022. Note the EC area, where there is an equal chance for above normal, normal, or below normal temperature conditions (white area between the 33 percent isolines).

Table 1 shows the comparison of skill scores for part of the temperature and precipitation data extending from the beginning of 2007 into 2008 (the full 182 month forecast comparison for the temperature and precipitation dataset is found in **Appendix 1**). The better skill score between WTI and CPC is color-coded for each forecast period. For the total period, WTI skill scores exceed the CPC scores for temperature in 155 of the 182 months, or 85.1 percent of the total forecasts issued during this time period. Thus, on average during this period, WTI bested the CPC on almost 6 out of every 7 forecasts, a highly statistically significant difference in forecast accuracy.

For precipitation, WTI skill scores exceeded those for CPC in 157 of 182 months, or 86.3 percent of the total forecasts issued during the time period. These results were slightly better than those for temperature.

Start Year	Months	WTI Skill Score (Temp)	CPC Skill Score (Temp)	WTI Skill Score (Precip)	CPC Skill Score (Precip)
	NDJ	60.19%	28.45%	19.23%	1.08%
	DJF	60.77%	14.44%	16.93%	-9.70%
2007	JFM	47.50%	21.12%	18.08%	-2.80%
	FMA	14.04%	6.90%	25.58%	-4.96%

Start Year	Months	WTI Skill Score (Temp)	CPC Skill Score (Temp)	WTI Skill Score (Precip)	CPC Skill Score (Precip)
	MAM	43.46%	32.33%	16.93%	10.34%
	AMJ	29.62%	0.00%	15.77%	1.29%
	MJJ	23.27%	0.43%	6.54%	1.08%
	JJA	20.97%	25.43%	9.43%	0.65%
	JAS	26.73%	45.47%	6.54%	-7.11%
	ASO	45.20%	35.99%	18.66%	-9.70%
	SON	48.66%	59.27%	-0.38%	3.88%
	OND	73.46%	40.30%	19.81%	12.93%
	NDJ	56.73%	15.09%	0.77%	20.04%
	DJF	42.89%	10.13%	28.47%	14.44%
2008	JFM	34.81%	-4.74%	13.47%	20.26%
	FMA	35.96%	-6.68%	17.50%	12.28%
	MAM	49.81%	-6.03%	7.70%	14.66%

Table 1. Temperature skill scores for a portion of the evaluation period. Skill scores are based upon the Heidke algorithm found above. For the skill score columns, light green colored skill scores designate the winning score for either WTI or CPC.
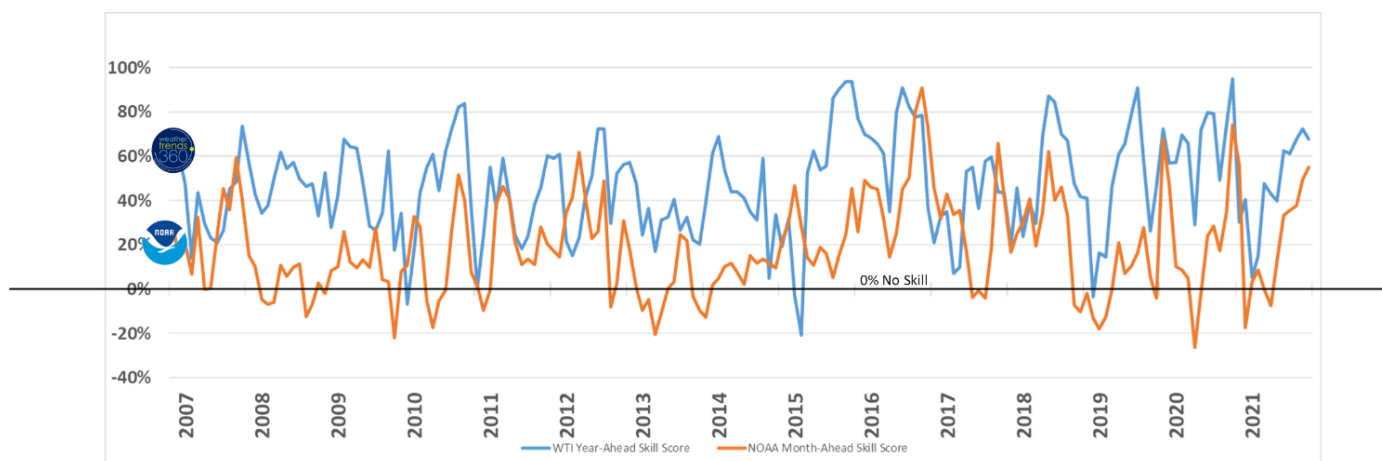
TEMPERATURE EVALUATION

In a vast majority of the three month forecast periods, WTI scored positive Heidke skill scores, indicating many more forecast "hits" than "misses". During the 182-month forecast period, WTI registered only 4 missed forecasts as defined by Heidke skill scores (Figure 2; bottom two quadrats), which represents a 97.8% percent accuracy rate based upon skill scores, all the more impressive considering that these forecasts were developed 11 months prior to the actual forecast date. CPC posted 36 missed forecasts (two quadrats to the left), which is an approximately 80.2 percent accuracy rate. It also represents approximately 9 times more missed forecasts than WTI, a significantly higher rate. The diagonal line in Figure 2 represents equal forecast accuracy by the two groups; dots to the left of that line represent a better skill score (forecast accuracy) by WTI; dots to the right represent a better accuracy by CPC. It is clear that the dots are strongly skewed to the left, indicating the superior skill scores by WTI. In fact, CPC had a superior skill score compared to WTI in just 27 of 182 forecast periods, a 14.8 percent rate (this can also be seen in the full dataset, Appendix 1). Thus, WTI's forecasts were more accurate than CPC's in approximately 6 out of every 7 forecasts during the evaluation period.
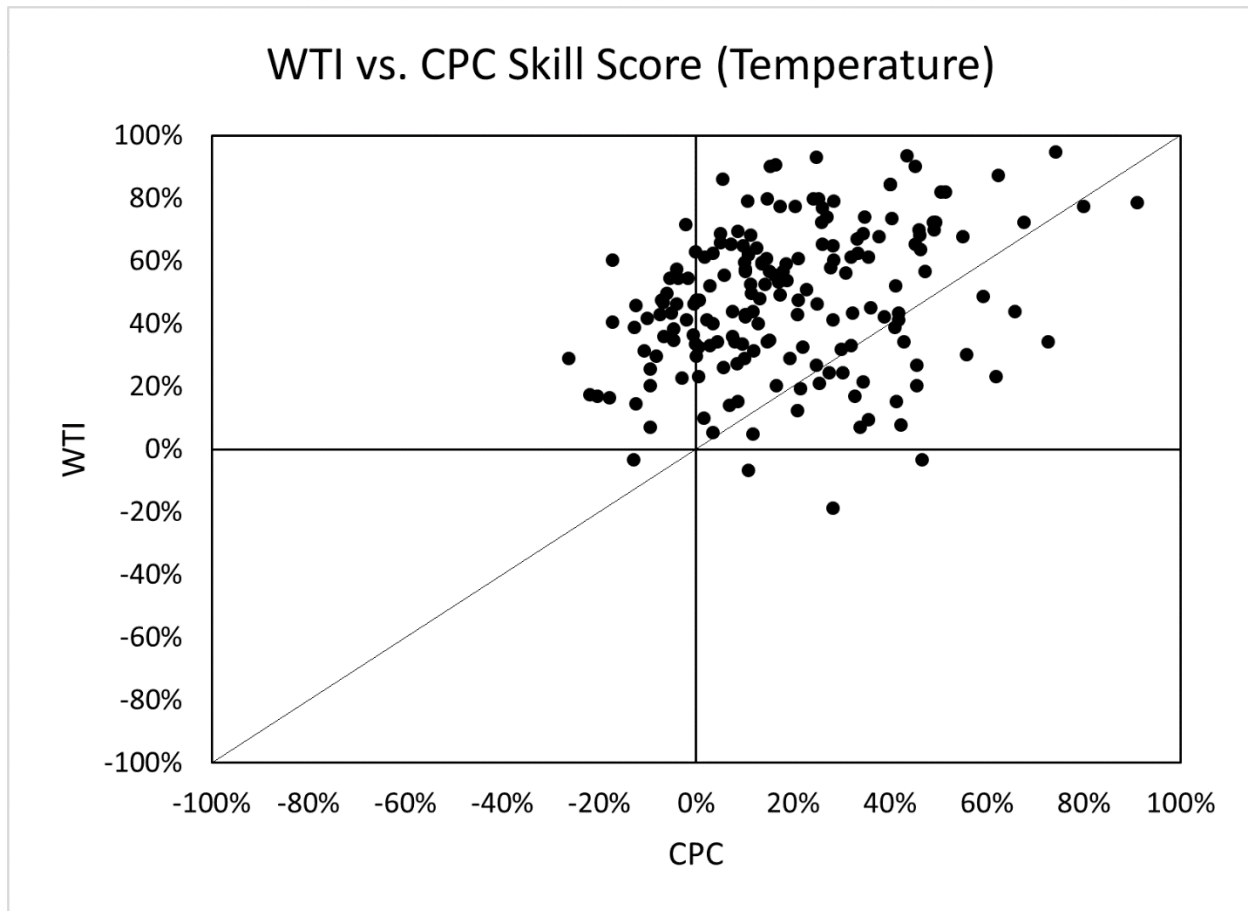
Figure 2. A comparison of temperature forecast "hits" vs. "misses" based upon Heidke skill scores for the 182-month evaluation period. The upper right quadrat includes forecast hits by both WTI and CPC. The upper left quadrat represents hits by WTI and misses by CPC. The lower right quadrat represents hits by CPC and misses by WTI. The lower left quadrat represents misses by both groups. Any points above the diagonal line indicate that WTI scores exceeded CPC's.

A temporal analysis of the temperature skill scores is instructive to see when forecast differentials were greatest (Figure 3). It is clear that WTI had greater forecast accuracy during a large majority of the time span, with the exception of some short periods scattered throughout the forecast period. At no point did CPC exceed WTI's skill score for more than four consecutive months, yet WTI's skill scores were higher for a 20 consecutive month period in 2008 and 2009, and a 28 consecutive month time span during 2012 to 2014.
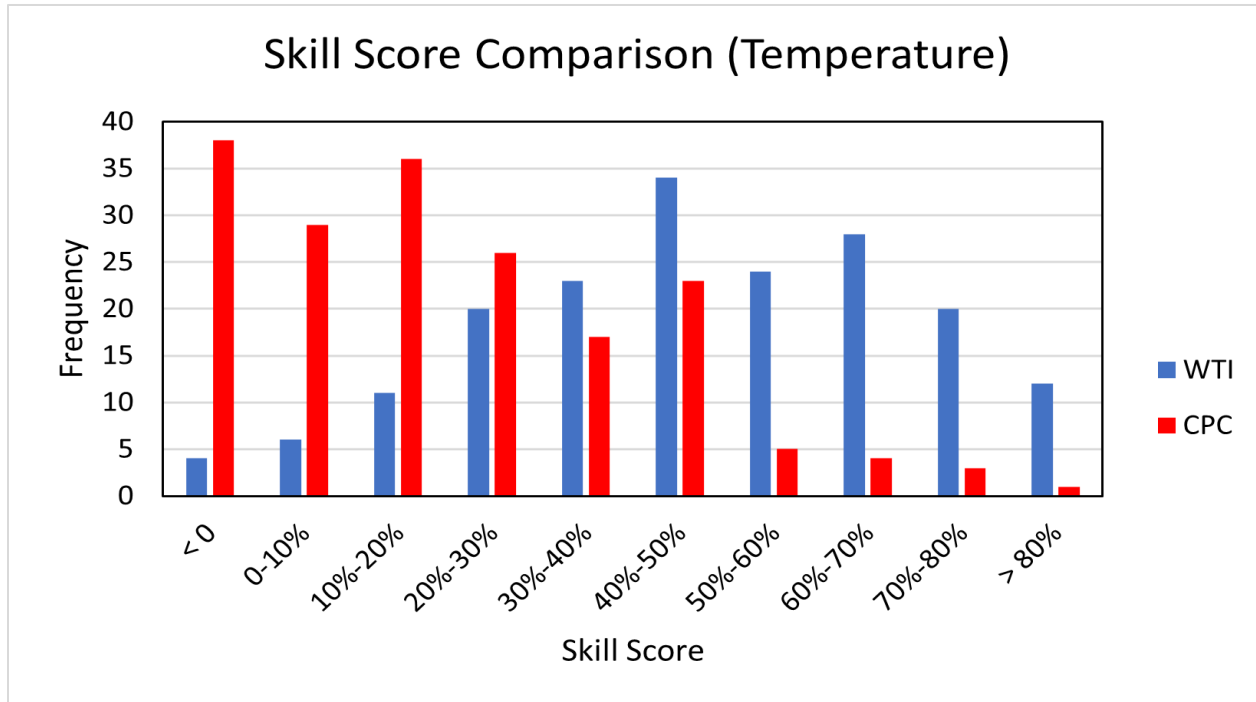
Figure 3. Skill score temperature comparison through the evaluation period. The difference between WTI's blue line and CPC orange line represent periods when WTI's accuracy exceeded CPC's.

We also evaluated the magnitude of the skill scores using a series of intervals (Figure 4). Once again, WTI shows a clearly superior set of scores, translating into more accurate forecasts. CPC's more negative scores, and slightly positive scores, stand in contrast to the larger number of highly positive scores that are indicative of the WTI forecasts. The switch occurs at the 30% Heidke skill score; all intervals higher than that show WTI forecasts as more dominant; lower intervals are more populated by CPC forecasts, and by a sizable margin. The mode for WTI scores is between 40-50%, considerably higher than CPC's mode.

The analysis strongly suggests that WTI has offered much more accurate forecasts than CPC over the period of evaluation. This is in spite of the fact that the WTI forecasts are issued 11 months prior to the actual forecast period, while CPC's forecasts are issued a month before. In summary, the data indicate that WTI issues a more accurate forecast over 85 percent of the time, CPC's forecasts are associated with negative skill score values at a rate 2.6 times higher than WTI's, and WTI forecasts are associated with higher skill scores than CPC's in 3 of every 4 months, on average, during the evaluation period.

There has been little change in CPC skill scores through the period of record (Figure 5). Through the 182-month evaluation period, the CPC skill scores have increased slightly, as shown by the trend line, however this increase is not statistically significant.

## Skill Score Comparison (Temperature)

Figure 4. A comparison of temperature skill score intervals, shown in tenths. Negative skill scores are to the far left; the most accurate skill scores are to the right.

CPC Skill Score (Temp)
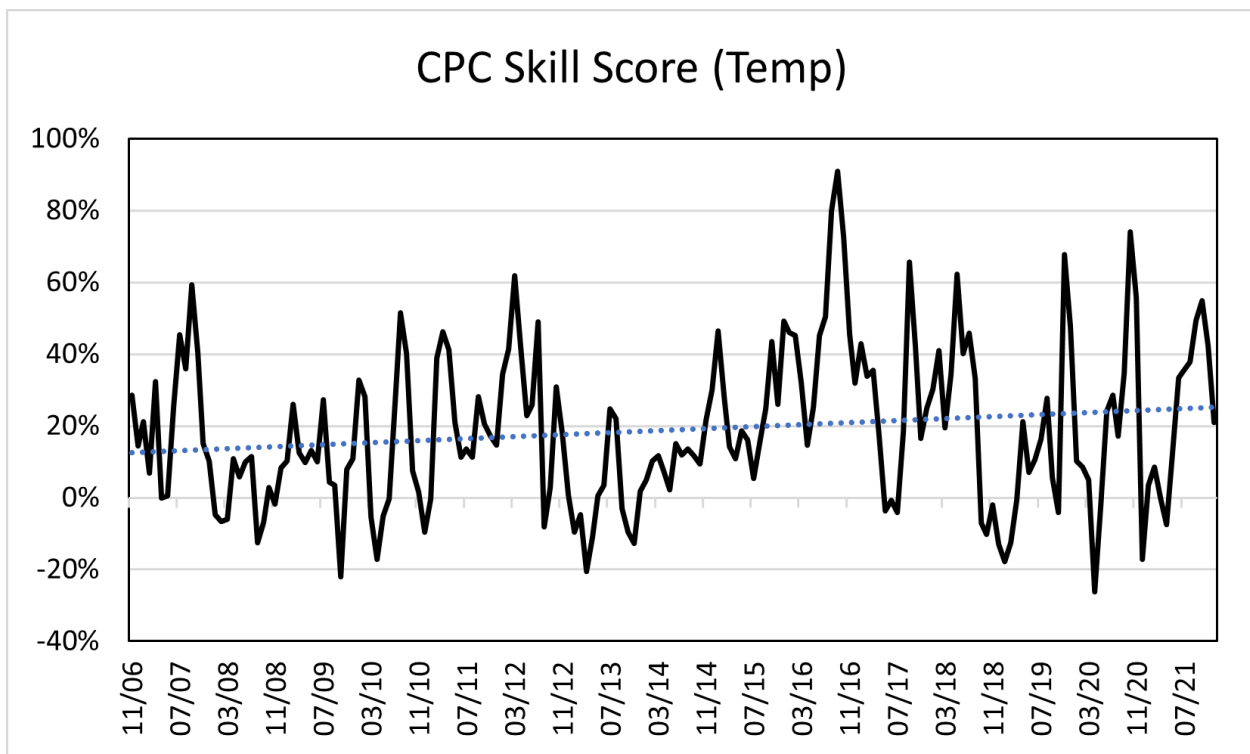
Figure 5. CPC temporal evaluation of temperature skill scores. Y-axis represents Heidke skill score value, x-axis represents month. Dotted line represents trend.

PRECIPITATION EVALUATION

There are many similarities between the temperature evaluation and the precipitation evaluation; in both, WTI skill scores exceed CPC skill scores by a significant amount (Figure 6).
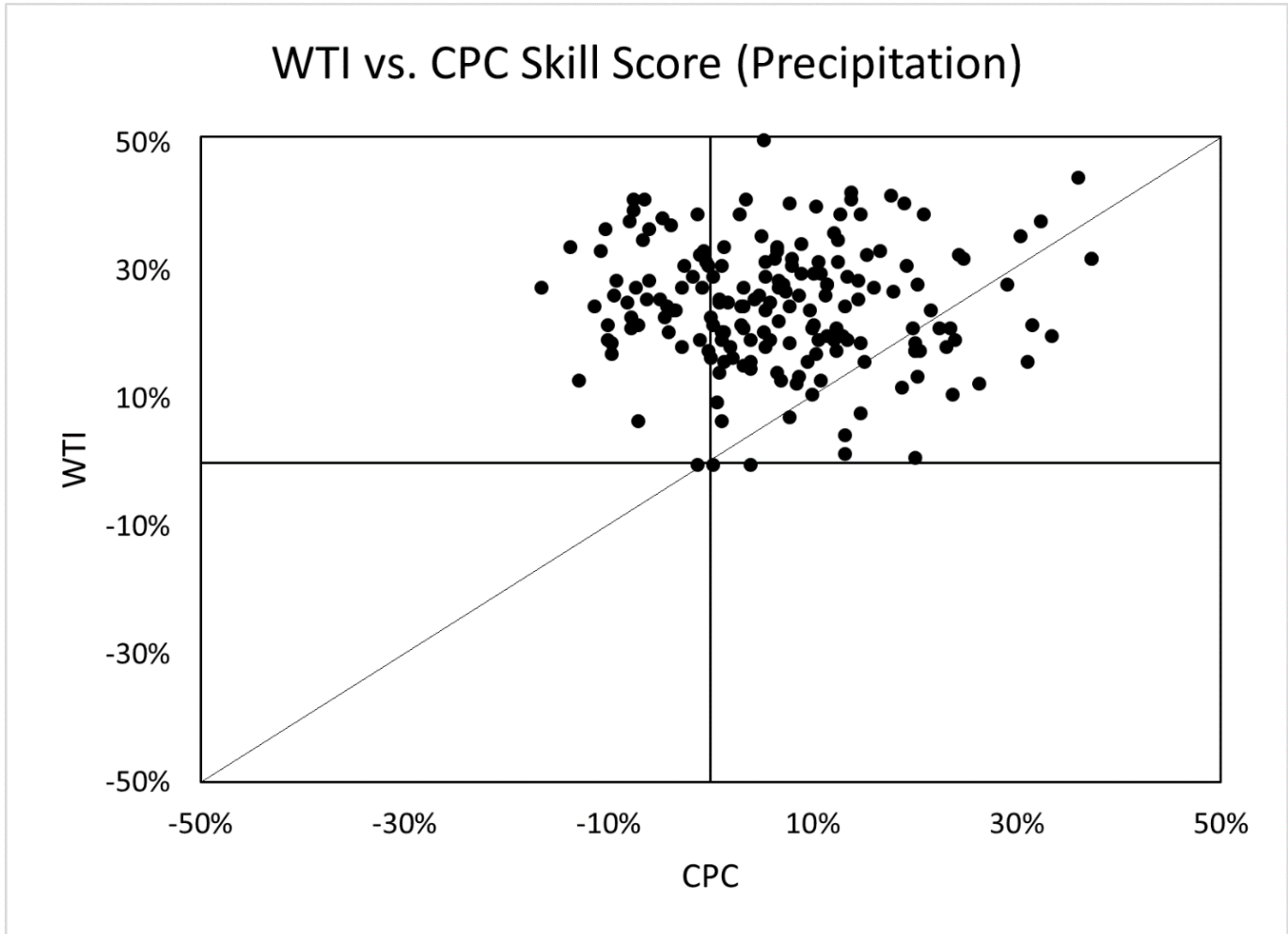
Figure 6. A comparison of precipitation forecast "hits" vs. "misses" based upon Heidke skill scores for the 182-month evaluation period. The upper right quadrat includes forecast hits by both WTI and CPC. The upper left quadrat represents hits by WTI and misses by CPC. The lower right quadrat represents hits by CPC and misses by WTI. The lower left quadrat represents misses by both groups. Any points above the diagonal line indicate that WTI scores exceeded CPC's.

Both CPC and WTI register "hits" for the majority of the time (upper right quadrat), but CPC shows considerably more "misses", as indicated by dots in the upper left quadrat. During the 182-month period, WTI shows only three misses (dots in two lower quadrats), representing an accuracy rate of 98.4%. CPC miss rates are considerably higher; there were 47 misses during the period, representing an accuracy rate of 74.2%.

Much like the temperature evaluation, the points are heavily skewed to the left of the diagonal line, indicating WTI's superior skill scores. WTI skill scores exceeded CPC's on 157 of the 182 3-month periods, which is 86.3% of the time. This result for WTI/CPC precipitation comparison is slightly better than the temperature results discussed earlier in the evaluation.

The temporal analysis of the precipitation skill scores shows when the forecast differentials are greatest (Figure 7). Clearly, the WTI skill scores are superior throughout the evaluation period, and this often occurs for long consecutive monthly periods; the most notable occur during much of 2009 and 2010, and also in 2013 and 2014. The temporal variation in CPC skill scores is clearly higher than for WTI, as apparent by the large swings from period to period in the CPC scores.
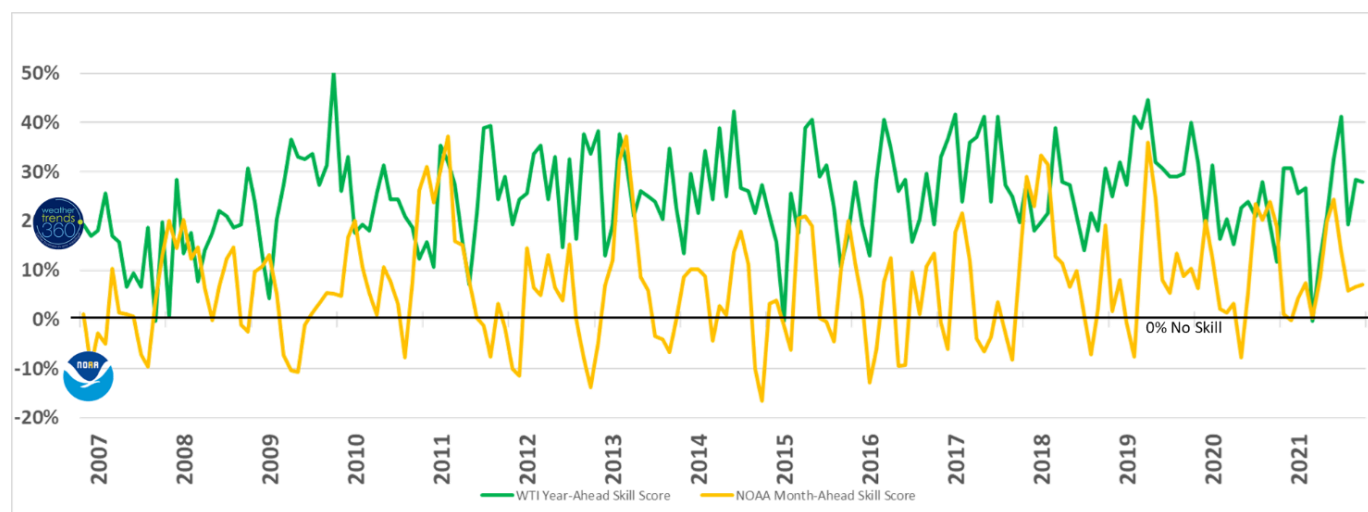
Figure 7. Skill score precipitation comparison through the evaluation period. The delta between WTI's skill (green line) and CPC (orange line) represent periods when WTI's accuracy exceeded CPC's in almost all cases.

We evaluated the magnitude of the precipitation skill scores using a series of intervals (Figure 8). Once again, WTI shows a clearly superior set of scores, translating into more accurate forecasts. CPC's more frequent negative scores, and frequent slightly positive scores, stand in contrast to the larger number of more highly positive scores that are indicative of the WTI forecasts. In general, skill scores for precipitation are lower for both CPC and WTI than those for temperature when comparing Figures 4 and 8. Many skill scores exceed 50% for temperature, while almost none reach that level for precipitation. As with the temperature histogram, the skill score mode is considerably higher for WTI when compared to CPC.
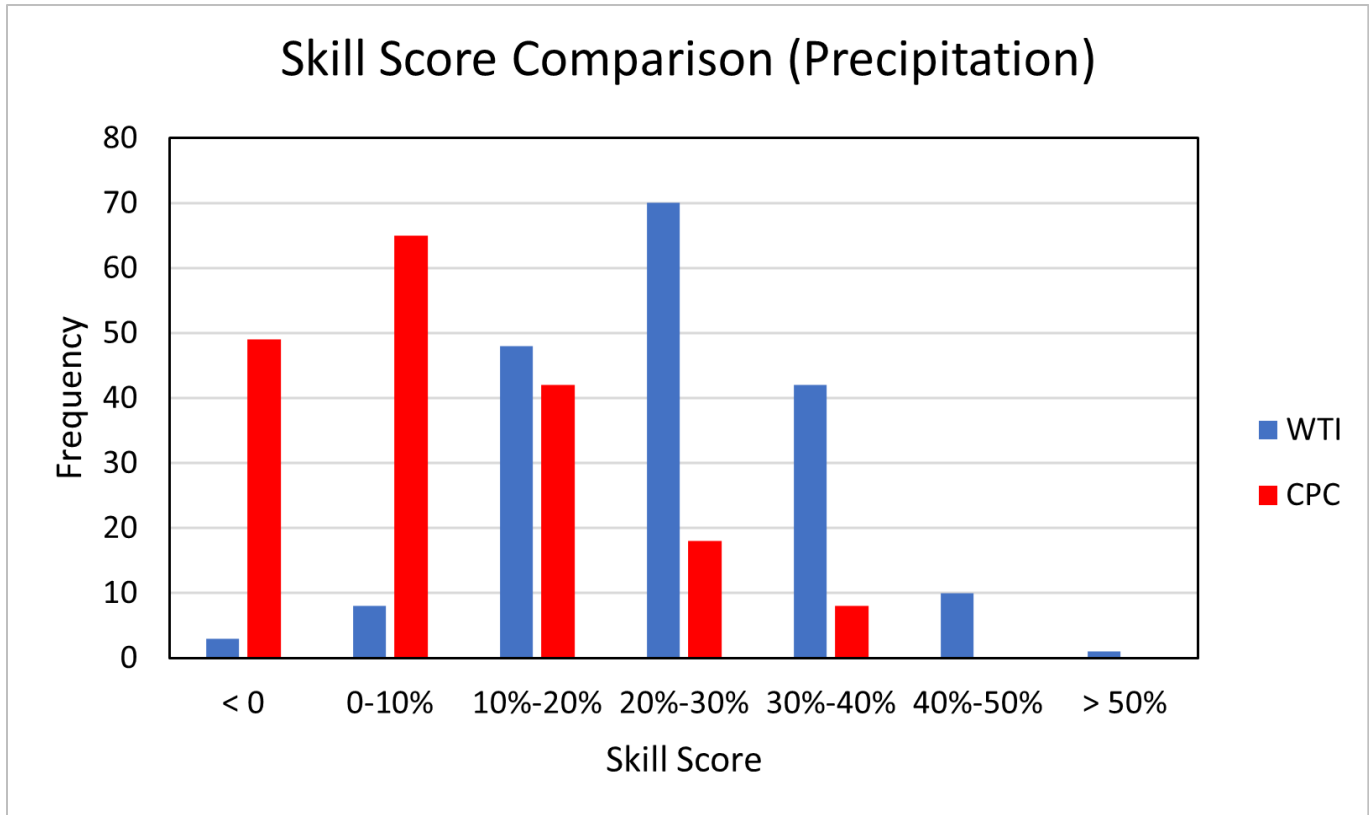
Figure 8. A comparison of precipitation skill score intervals, shown in tenths. Negative skill scores are to the left; the most accurate skill scores are to the right.

During the evaluated time period, CPC skill scores have increases slightly, which is a positive development (Figure 9). However, this increase is not nearly sufficient to reach accuracy levels achieved by WTI in forecasting precipitation. The higher CPC temporal variability in forecasting also stands in contrast with WTI's much more consistent positive skill scores.

The evaluation strongly suggests that WTI has produced consistently better temperature and precipitation forecasting than has CPC during the evaluation period. By every metric used here there is a much higher probability that WTI will produce a more accurate forecast than will CPC into the future.
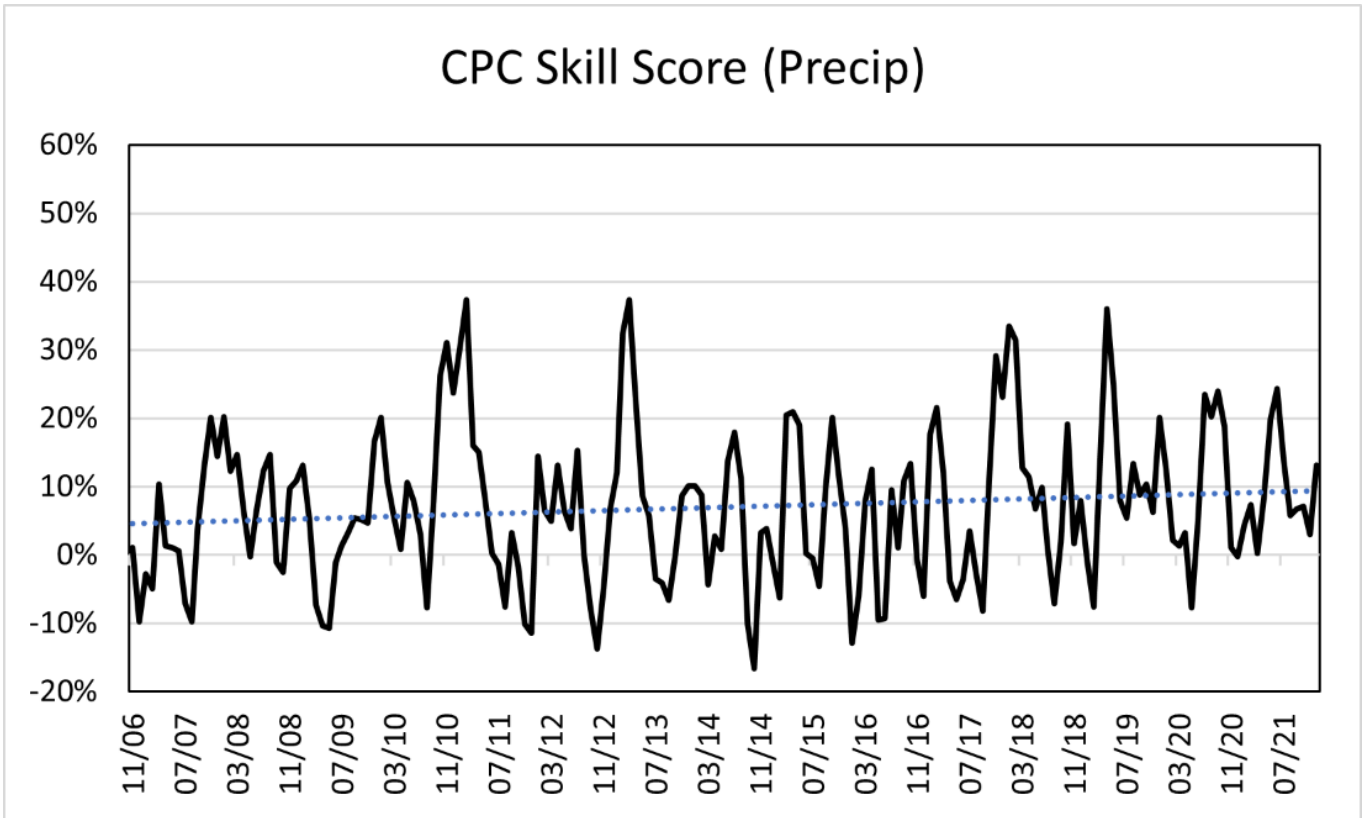
Figure 9. CPC temporal evaluation of precipitation skill scores. Y-axis represents Heidke skill score value, x-axis represents month. Dotted line represents trend.

APPENDIX

SKILL SCORE VALUES FOR THE ENTIRE EVALUATION PERIOD

This is the total dataset for the entire evaluation period constructed by WTI. Skill scores for temperature are in the left colored columns; those for precipitation are on the right. Colored scores represent the higher skill score for either WTI or CPC.

Start Year	Months	WTI Skill Score (Temp)	CPC Skill Score (Temp)	WTI Skill Score (Precip)	CPC Skill Score (Precip)
2006	NDJ	60.19%	28.45%	19.23%	1.08%
	DJF	60.77%	14.44%	16.93%	-9.70%
2007	JFM	47.50%	21.12%	18.08%	-2.80%
	FMA	14.04%	6.90%	25.58%	-4.96%
	MAM	43.46%	32.33%	16.93%	10.34%
	AMJ	29.62%	0.00%	15.77%	1.29%
	MJJ	23.27%	0.43%	6.54%	1.08%
	JJA	20.97%	25.43%	9.43%	0.65%
	JAS	26.73%	45.47%	6.54%	-7.11%
	ASO	45.20%	35.99%	18.66%	-9.70%
	SON	48.66%	59.27%	-0.38%	3.88%
	OND	73.46%	40.30%	19.81%	12.93%
	NDJ	56.73%	15.09%	0.77%	20.04%
	DJF	42.89%	10.13%	28.47%	14.44%
2008	JFM	34.81%	-4.74%	13.47%	20.26%
	FMA	35.96%	-6.68%	17.50%	12.28%
	MAM	49.81%	-6.03%	7.70%	14.66%
	AMJ	61.92%	10.78%	14.04%	6.47%
	MJJ	55.58%	5.82%	17.50%	-0.22%
	JJA	59.62%	9.91%	22.12%	6.68%
	JAS	49.81%	11.42%	20.97%	12.28%
	ASO	45.77%	-12.50%	18.66%	14.66%
	SON	46.92%	-6.90%	19.23%	-1.08%
	OND	33.08%	2.80%	30.77%	-2.59%
	NDJ	54.42%	-1.72%	23.85%	9.70%
	DJF	27.31%	8.41%	12.89%	10.78%
2009	JFM	42.31%	10.13%	4.24%	13.15%
	FMA	65.38%	26.08%	20.39%	5.17%
	MAM	64.23%	12.50%	27.31%	-7.33%
	AMJ	64.81%	9.70%	36.54%	-10.34%
	MJJ	48.08%	13.15%	33.08%	-10.78%
	JJA	29.04%	9.91%	32.50%	-1.08%
	JAS	24.42%	27.37%	33.66%	1.29%

Start Year	Months	WTI Skill Score (Temp)	CPC Skill Score (Temp)	WTI Skill Score (Precip)	CPC Skill Score (Precip)
	ASO	34.23%	4.31%	27.31%	3.23%
	SON	62.50%	3.45%	31.35%	5.39%
	OND	17.50%	-21.98%	50.39%	5.17%
	NDJ	34.23%	7.97%	26.16%	4.74%
	DJF	-6.73%	10.78%	33.08%	16.59%
2010	JFM	16.92%	32.76%	17.50%	20.04%
	FMA	41.15%	28.23%	19.23%	10.56%
	MAM	54.42%	-5.39%	18.08%	5.39%
	AMJ	60.19%	-17.24%	25.58%	0.86%
	MJJ	43.46%	-5.17%	31.35%	10.56%
	JJA	63.08%	-0.22%	24.43%	7.76%
	JAS	74.04%	26.94%	24.43%	3.02%
	ASO	82.12%	51.51%	20.97%	-7.76%
	SON	84.42%	40.09%	18.66%	7.76%
	OND	35.96%	7.54%	12.31%	26.29%
	NDJ	10.00%	1.51%	15.77%	31.03%
	DJF	7.12%	-9.48%	10.58%	23.71%
2011	JFM	33.65%	-0.22%	35.39%	30.39%
	FMA	42.31%	38.79%	31.93%	37.28%
	MAM	63.65%	46.34%	27.31%	15.95%
	AMJ	52.12%	41.16%	15.77%	15.09%
	MJJ	42.88%	20.91%	7.12%	7.76%
	JJA	52.69%	11.21%	21.54%	0.22%
	JAS	59.62%	13.58%	38.85%	-1.29%
	ASO	68.27%	11.21%	39.43%	-7.54%
	SON	64.81%	28.23%	24.43%	3.23%
	OND	77.50%	20.47%	29.04%	-1.72%
	NDJ	77.50%	17.24%	19.23%	-10.13%
	DJF	79.81%	14.66%	24.43%	-11.42%
2012	JFM	21.54%	34.48%	25.58%	14.44%
	FMA	15.19%	41.38%	33.66%	6.47%
	MAM	23.27%	61.85%	35.39%	4.96%
	AMJ	41.15%	41.81%	24.43%	13.15%
	MJJ	50.96%	22.84%	33.08%	6.47%
	JJA	72.31%	25.86%	14.62%	3.88%
	JAS	72.31%	48.92%	32.50%	15.30%
	ASO	29.62%	-8.19%	16.35%	0.00%
	SON	52.12%	2.80%	37.70%	-7.97%
	OND	56.15%	30.82%	33.66%	-13.79%
	NDJ	56.73%	17.89%	38.27%	-4.74%
	DJF	47.50%	0.65%	12.89%	6.90%
2013	JFM	25.58%	-9.48%	19.23%	12.07%
	FMA	38.27%	-4.74%	37.70%	32.33%

Start Year	Months	WTI Skill Score (Temp)	CPC Skill Score (Temp)	WTI Skill Score (Precip)	CPC Skill Score (Precip)
	MAM	16.92%	-20.47%	31.93%	37.28%
	AMJ	31.35%	-10.78%	20.97%	22.41%
	MJJ	32.50%	0.43%	26.16%	8.62%
	JJA	40.00%	3.45%	25.00%	5.82%
	JAS	26.73%	24.78%	23.85%	-3.45%
	ASO	32.50%	21.98%	20.39%	-4.09%
	SON	22.69%	-3.02%	34.81%	-6.68%
	OND	20.38%	-9.48%	22.70%	0.00%
	NDJ	38.85%	-12.72%	13.47%	8.62%
	DJF	61.35%	1.72%	29.62%	10.13%
2014	JFM	68.85%	4.96%	21.54%	10.13%
	FMA	56.73%	10.13%	34.23%	8.84%
	MAM	44.04%	11.64%	24.43%	-4.31%
	AMJ	44.04%	7.54%	38.85%	2.80%
	MJJ	41.15%	2.16%	25.00%	0.86%
	JJA	34.81%	15.09%	42.31%	13.79%
	JAS	31.35%	11.85%	26.73%	17.89%
	ASO	59.04%	13.58%	26.16%	11.21%
	SON	4.81%	11.64%	21.54%	-10.13%
	OND	33.65%	9.48%	27.31%	-16.59%
	NDJ	19.23%	21.55%	20.97%	3.23%
	DJF	31.92%	29.96%	15.77%	3.88%
2015	JFM	-3.27%	46.55%	-0.38%	-1.29%
	FMA	-18.85%	28.23%	25.58%	-6.25%
	MAM	52.69%	14.22%	17.50%	20.47%
	AMJ	62.50%	10.78%	38.85%	20.91%
	MJJ	53.85%	18.75%	40.58%	18.97%
	JJA	55.58%	16.16%	29.04%	0.22%
	JAS	86.15%	5.39%	31.35%	-0.43%
	ASO	90.19%	15.30%	22.70%	-4.53%
	SON	93.08%	24.78%	10.58%	9.91%
	OND	93.65%	43.53%	17.50%	20.04%
	NDJ	76.92%	26.08%	27.89%	11.42%
	DJF	70.00%	49.14%	19.23%	3.88%
2016	JFM	68.27%	46.12%	12.89%	-12.93%
	FMA	65.38%	45.26%	28.47%	-6.03%
	MAM	61.35%	31.90%	40.58%	7.76%
	AMJ	34.23%	14.66%	34.81%	12.50%
	MJJ	79.81%	25.22%	26.16%	-9.48%
	JJA	90.19%	45.26%	28.47%	-9.27%
	JAS	82.12%	50.43%	15.77%	9.48%
	ASO	77.50%	79.96%	20.39%	1.08%
	SON	78.65%	90.95%	29.62%	10.78%

Start Year	Months	WTI Skill Score (Temp)	CPC Skill Score (Temp)	WTI Skill Score (Precip)	CPC Skill Score (Precip)
	OND	34.23%	72.63%	19.23%	13.36%
	NDJ	20.38%	45.47%	33.08%	-0.65%
	DJF	33.08%	31.90%	36.54%	-6.03%
2017	JFM	34.23%	42.89%	41.73%	17.67%
	FMA	7.12%	33.84%	23.85%	21.55%
	MAM	9.42%	35.56%	35.96%	12.07%
	AMJ	53.27%	17.03%	37.12%	-3.88%
	MJJ	54.42%	-3.66%	41.16%	-6.47%
	JJA	36.54%	-0.65%	23.85%	-3.66%
	JAS	57.31%	-4.09%	41.16%	3.45%
	ASO	59.04%	18.53%	27.31%	-2.80%
	SON	44.04%	65.73%	25.00%	-8.19%
	OND	43.46%	41.81%	19.81%	11.42%
	NDJ	20.38%	16.59%	27.89%	29.09%
	DJF	46.35%	25.00%	18.08%	23.06%
2018	JFM	24.42%	30.17%	19.81%	33.41%
	FMA	38.85%	40.95%	21.54%	31.47%
	MAM	29.04%	19.40%	38.85%	12.72%
	AMJ	68.85%	34.48%	27.89%	11.42%
	MJJ	87.31%	62.28%	27.31%	6.68%
	JJA	84.42%	40.09%	20.97%	9.91%
	JAS	70.00%	45.91%	14.04%	0.86%
	ASO	67.12%	33.19%	21.54%	-7.11%
	SON	47.50%	-7.11%	18.08%	1.94%
	OND	41.73%	-10.13%	30.77%	19.18%
	NDJ	41.16%	-1.94%	25.00%	1.72%
	DJF	-3.26%	-12.93%	31.93%	7.97%
2019	JFM	16.35%	-17.89%	27.31%	-0.86%
	FMA	14.62%	-12.50%	41.16%	-7.54%
	MAM	46.35%	-0.43%	38.85%	14.66%
	AMJ	60.77%	21.12%	44.62%	35.99%
	MJJ	65.39%	7.11%	31.93%	24.78%
	JJA	79.23%	10.56%	30.77%	7.97%
	JAS	90.77%	16.38%	29.04%	5.39%
	ASO	57.89%	27.80%	29.04%	13.36%
	SON	26.16%	5.60%	29.62%	8.84%
	OND	46.35%	-4.09%	40.00%	10.34%
	NDJ	72.31%	67.67%	31.93%	6.25%
	DJF	56.73%	47.20%	18.66%	20.04%
2020	JFM	57.31%	10.13%	31.35%	12.50%
	FMA	69.42%	8.62%	16.35%	2.16%
	MAM	65.96%	4.96%	20.39%	1.29%
	AMJ	29.04%	-26.29%	15.20%	3.23%

Start Year	Months	WTI Skill Score (Temp)	CPC Skill Score (Temp)	WTI Skill Score (Precip)	CPC Skill Score (Precip)
	MJJ	71.73%	-2.16%	22.70%	-7.76%
	JJA	79.81%	24.14%	23.85%	5.39%
	JAS	79.23%	28.45%	20.97%	23.49%
	ASO	49.23%	17.24%	27.89%	20.26%
	SON	74.04%	34.70%	19.23%	23.92%
	OND	94.81%	74.14%	11.74%	18.75%
	NDJ	30.20%	55.82%	30.77%	1.08%
	DJF	40.58%	-17.24%	30.77%	-0.22%
2021	JFM	5.39%	3.45%	25.58%	4.31%
	FMA	15.20%	8.62%	26.73%	7.33%
	MAM	47.50%	0.00%	-0.38%	0.22%
	AMJ	42.89%	-7.54%	12.31%	8.41%
	MJJ	40.00%	12.72%	20.97%	19.83%
	JJA	62.50%	33.41%	32.50%	24.35%
	JAS	61.35%	35.56%	41.16%	13.79%
	ASO	67.69%	37.72%	19.23%	5.82%
	SON	72.31%	49.35%	28.47%	6.68%
	OND	67.69%	54.96%	27.89%	7.11%