NWRFC Verification Study

Joe Intermill
Senior Hydrologist - NWRFC

NOAA NWS ~ NORTHWEST RIVER FORECAST CENTER

NWRFC – Verification Case Study

- Verification Goals/Approach
- Quick Basin Overview
- Event Description
- IVP Results/Conclusions
- IFP Results/Conclusions
- Case Study Conclusions

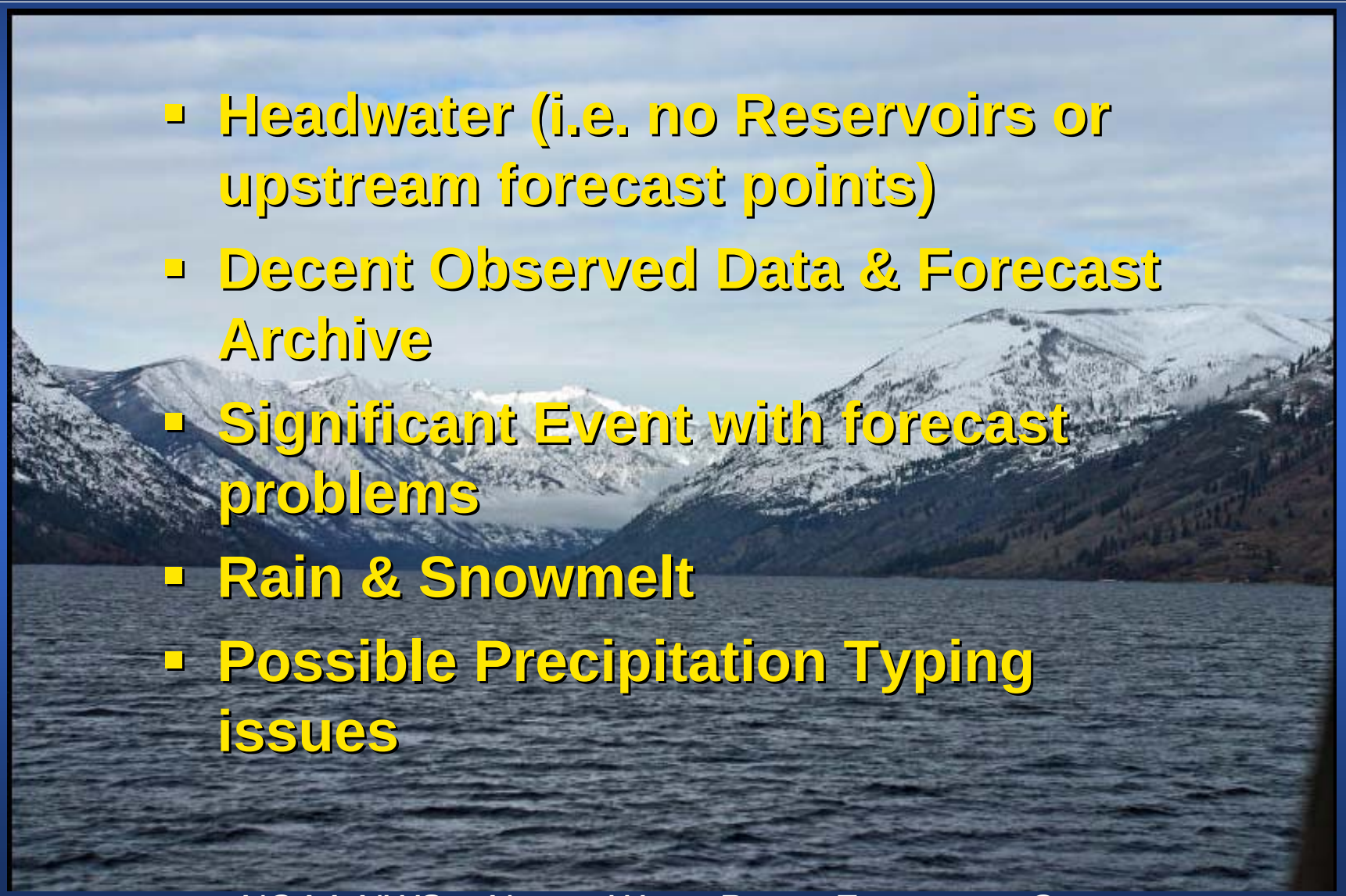
Verification Study Goals

- Assess validity and usefulness of existing Verification and re-analysis tools (i.e. IVP, RAX, RAX Apps and IFP) in the context of a single hydrologic event.
- Develop some conclusions/recommendations regarding the use and improvement of the tools.
- Develop a plan for systematic testing and implementation of these verification tools and techniques in a “real-time” environment.
- Use the information to improve our forecasts and understand their accuracy.

Verification Approach

- Use the Interactive Verification Program (IVP) and Archive Database (RAX) to analyze the event.
- The focus will be on Stage, Flow, MAP/FMAP, FMAT/MAT
- Use Archived versions of NWSRFS OFS files and the Interactive Forecast Program (IFP) as tools to disaggregate and analyze forecast components.
- Combine the results of the two analysis techniques and offer some conclusions on event verification.
- Make some comments on the apparent validity and usefulness of the various tools in the verification process. Can these tools help us to improve our forecasts?

Basin / Event Selection Process



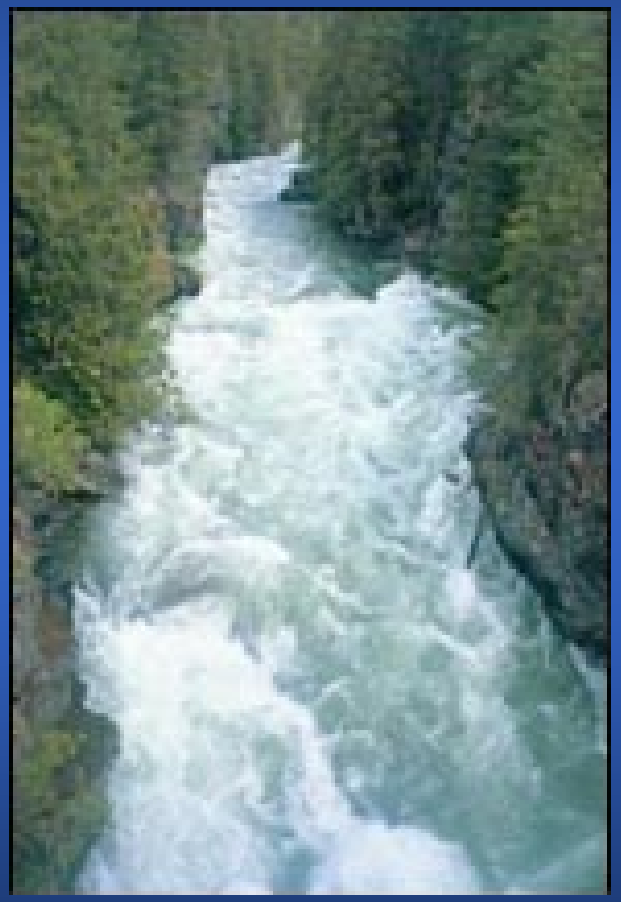
- Headwater (i.e. no Reservoirs or upstream forecast points)
- Decent Observed Data & Forecast Archive
- Significant Event with forecast problems
- Rain & Snowmelt
- Possible Precipitation Typing issues

Stehekin River at Stehekin

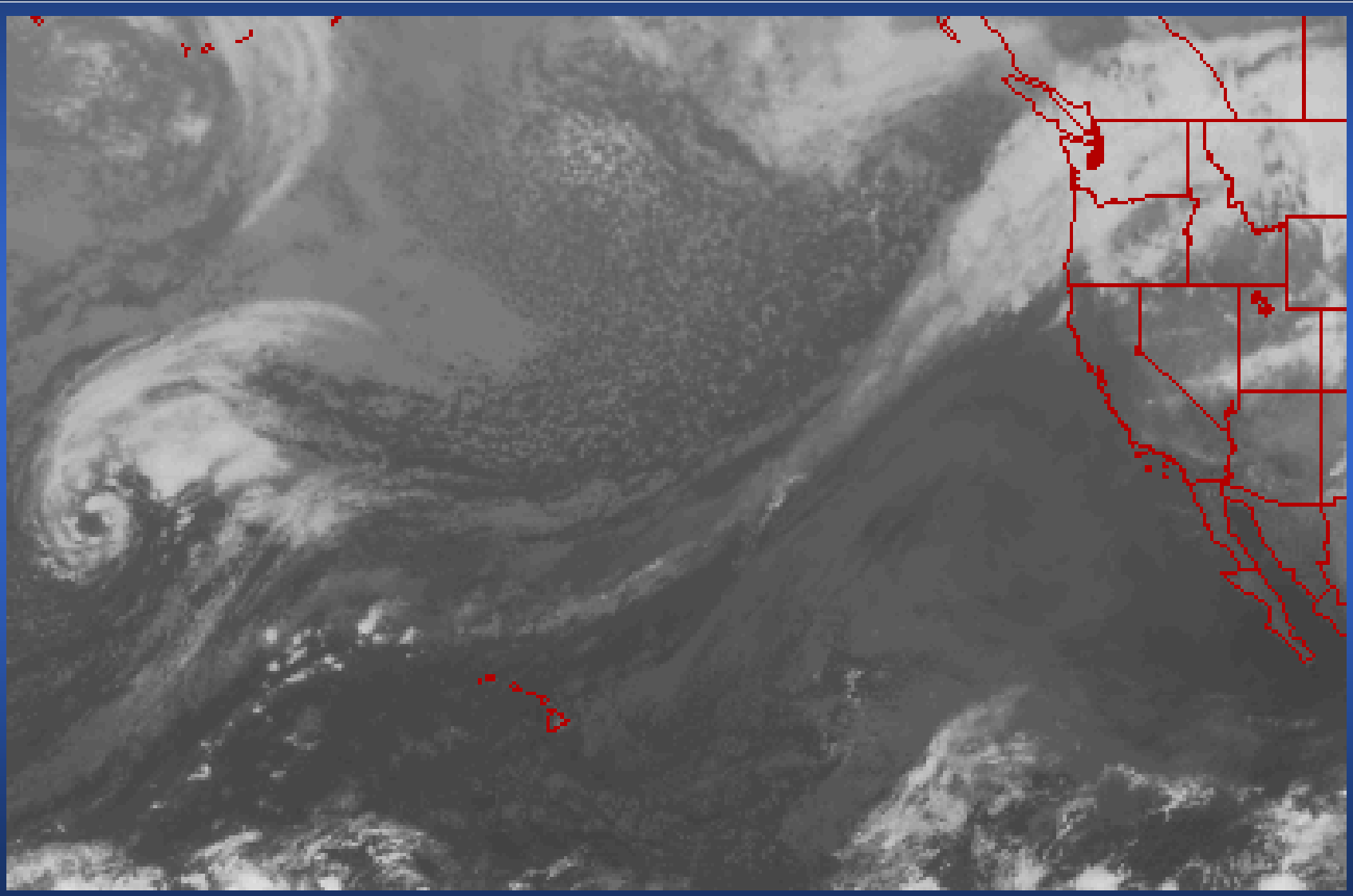
- North Central Washington
- Drains East Cascades
- Flows into Lake Chelan
- Significant rainfall and snowmelt runoff
- ~ 80% Forest Cover
- Basin Area - 321 sq. mi.
- Elev. Range: 1099 – 8760’
- Basin Mean Elev. ~ 5100’
- Glacier Cover ~ 2-3%
- Annual Precip ~ 80”

Event Description – Nov. 2006

- 6 Days of Widespread Heavy Precipitation in Westside basins spilled over to East Cascades (10-20 inch reports were common)
- Snow pack was **not** abnormally large in the basin (less than 5 inches of SWE at most area SNOTEL sites prior to heaviest precipitation).
- Temperatures early in the event were near rain/snow threshold for much of the basin. During the heaviest periods, precipitation fell as rain throughout the basin.
- Soil Moisture and river level conditions were normal-to-below normal leading up to the event.

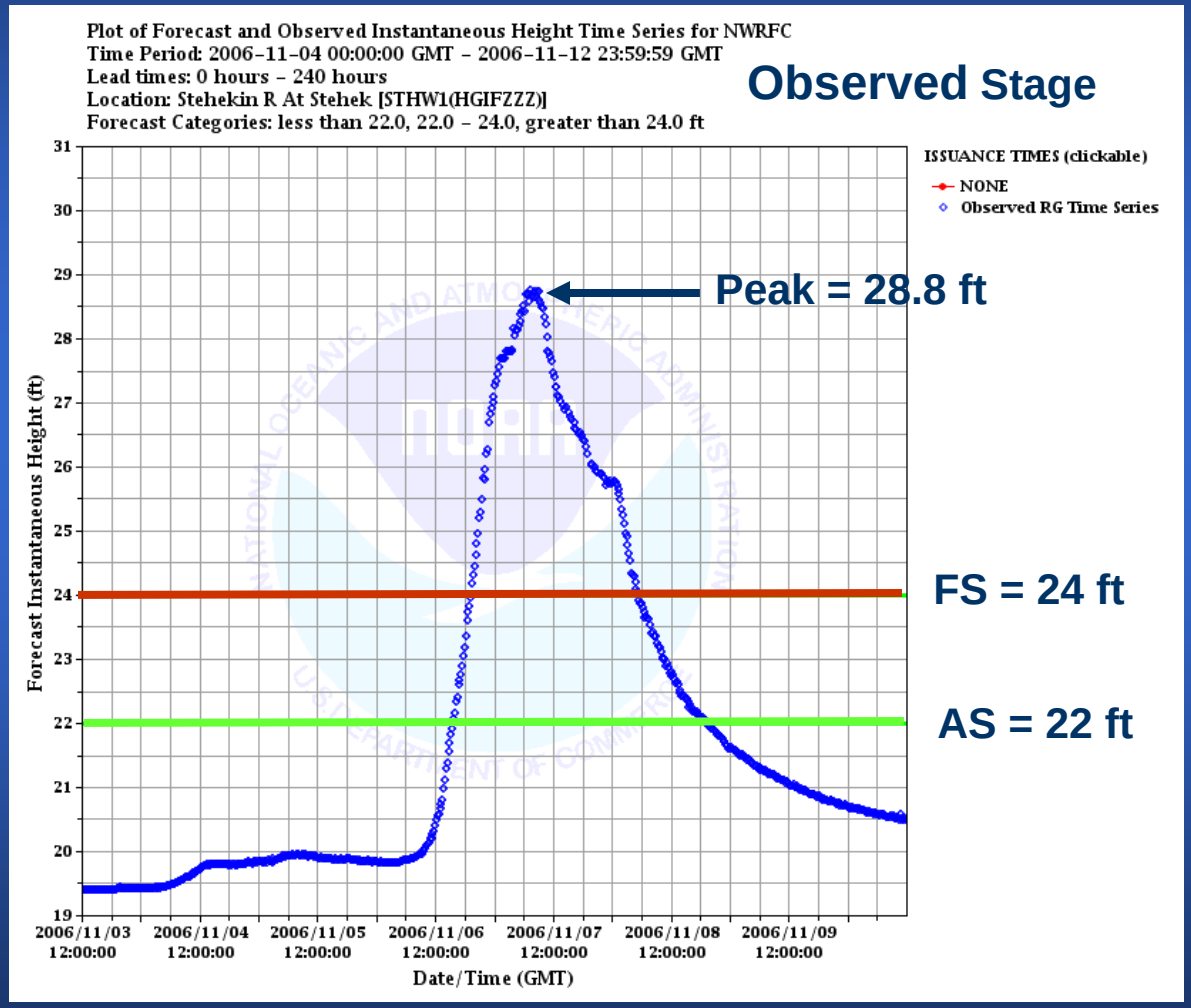


GOES IR - Nov 7, 2006 12Z



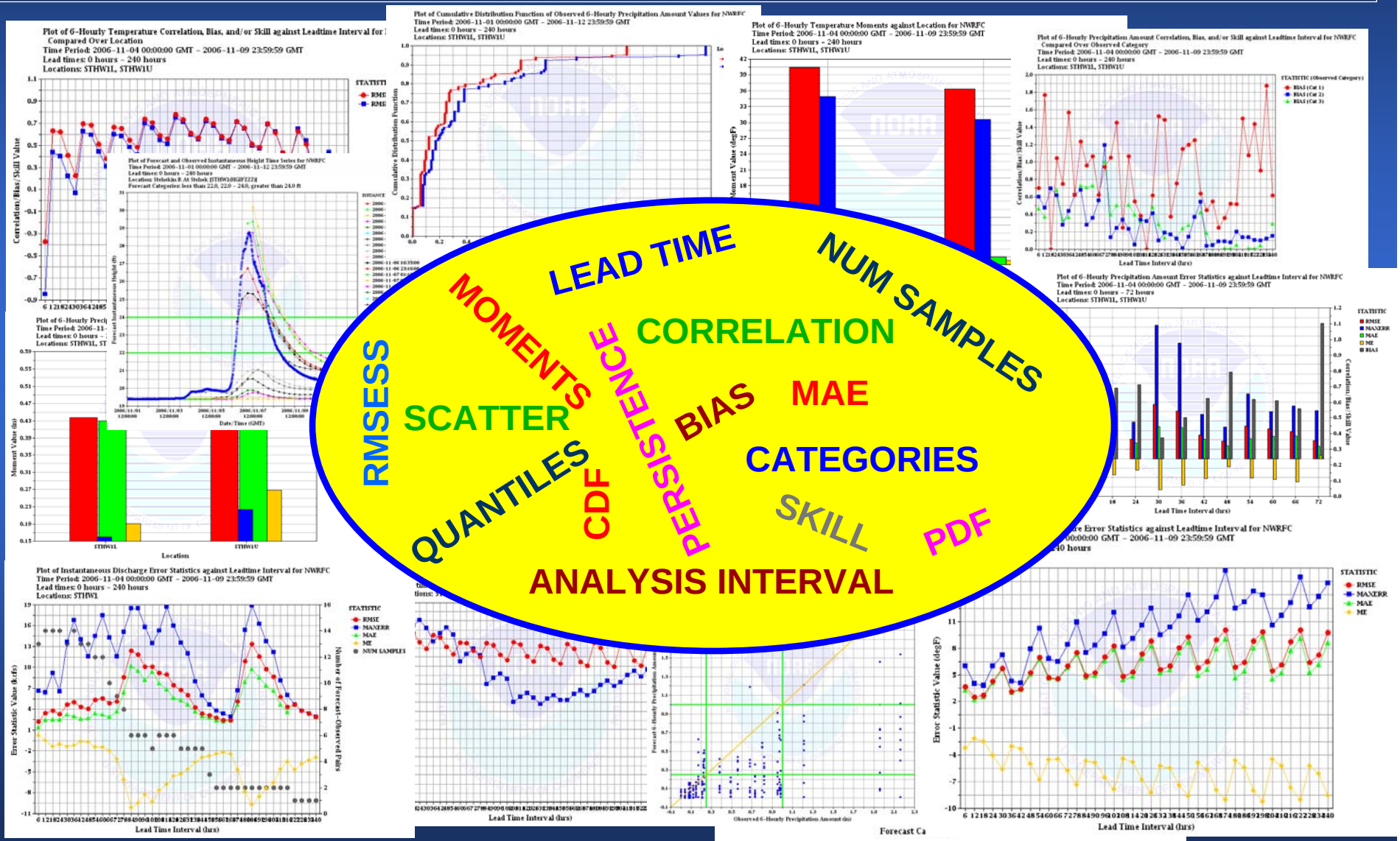
NOAA NWS ~ NORTHWEST RIVER FORECAST CENTER

Stehekin River Hydrograph: Nov. 4-9, 2006

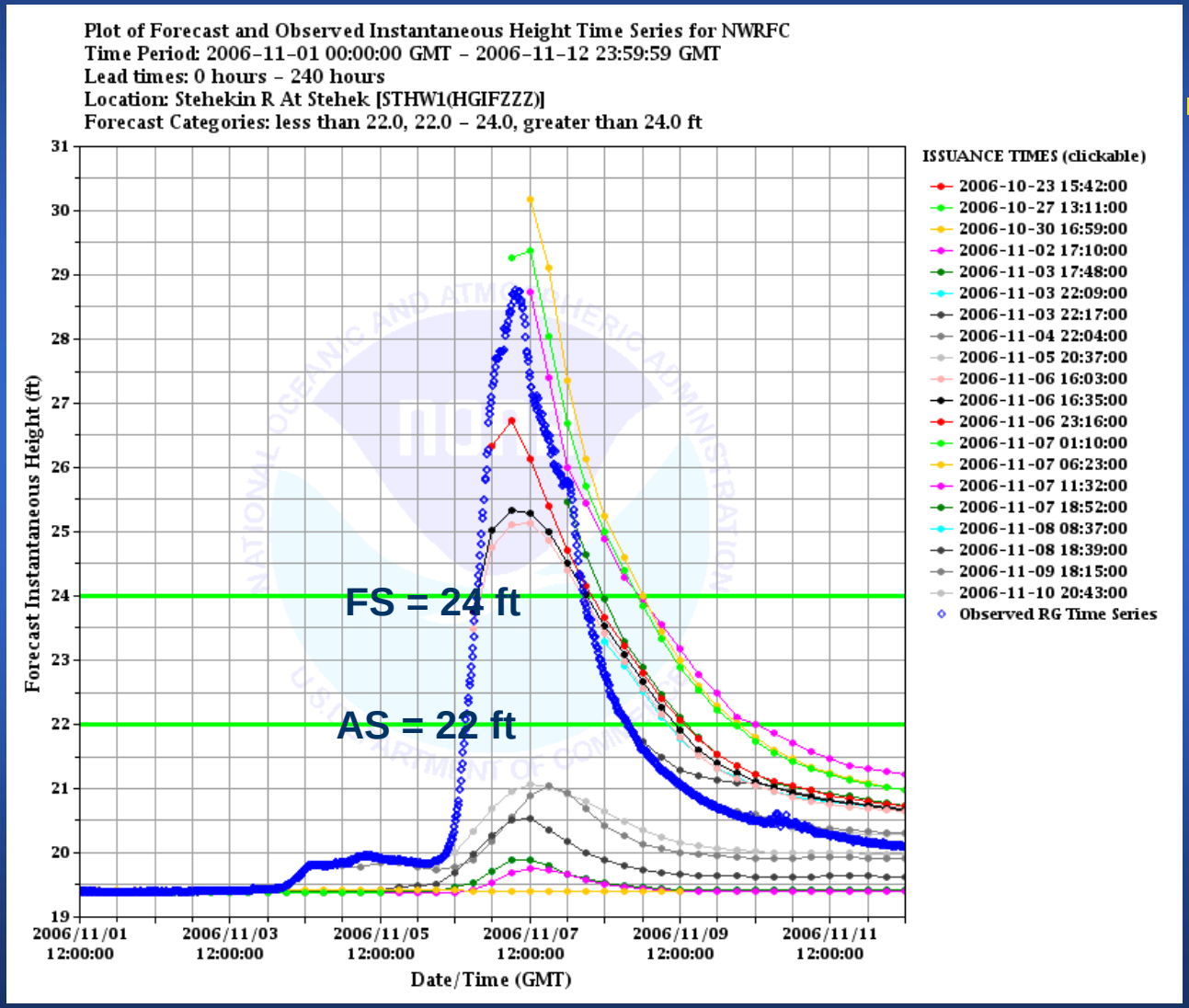


- Background
  - Event Peak = 28.8 ft. (11/07/2006 ~08Z)
  - Flood Stage = 24 ft
  - Action Stage = 22 ft

IVP ANALYSIS



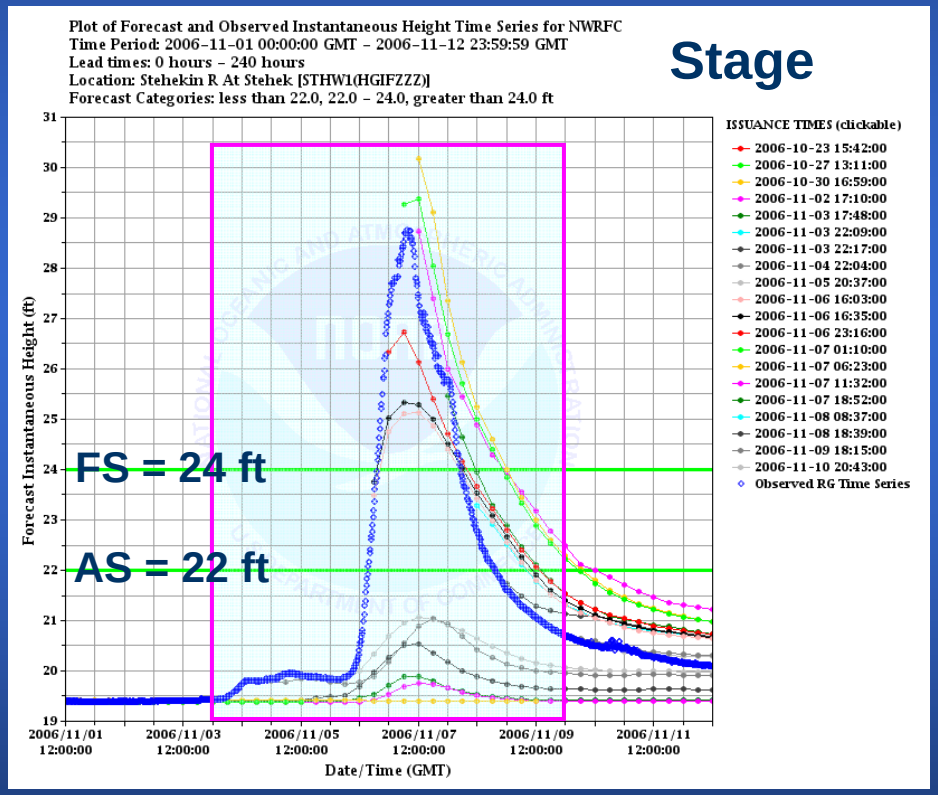
Time Series Plot for: Observed and Forecast Stage (STHW1)

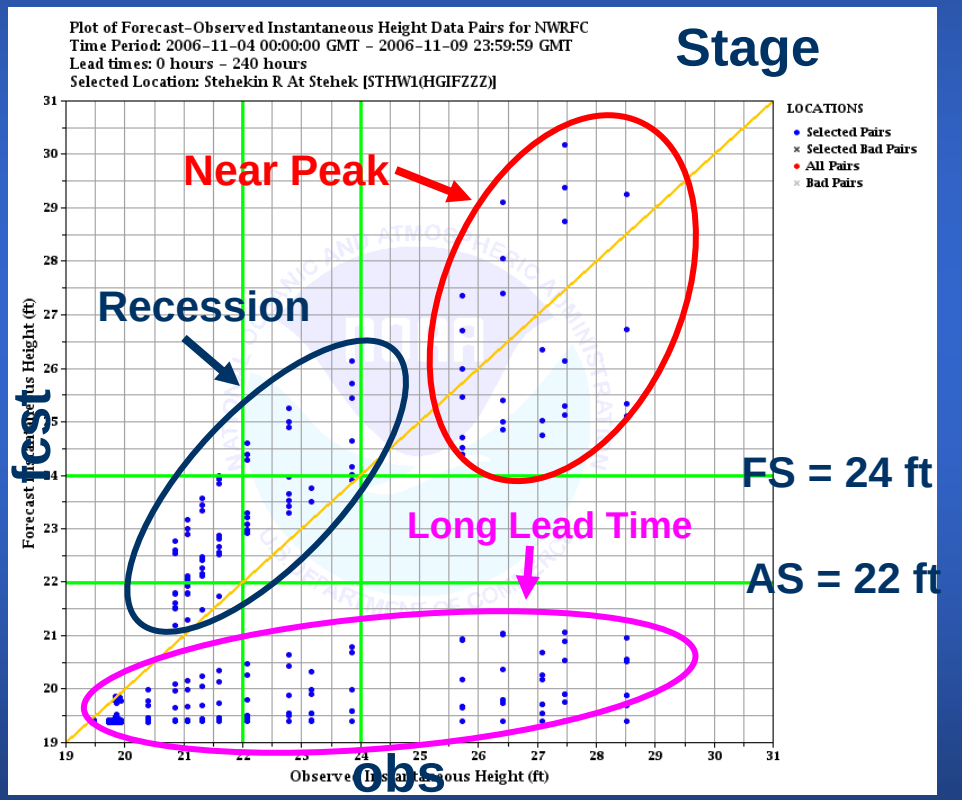


- IVP Parameters
  - Analysis Period: Nov. 4 - 9
  - Look at forecasts from 0 hrs to 10 day lead time
  - 6-hour Time Step

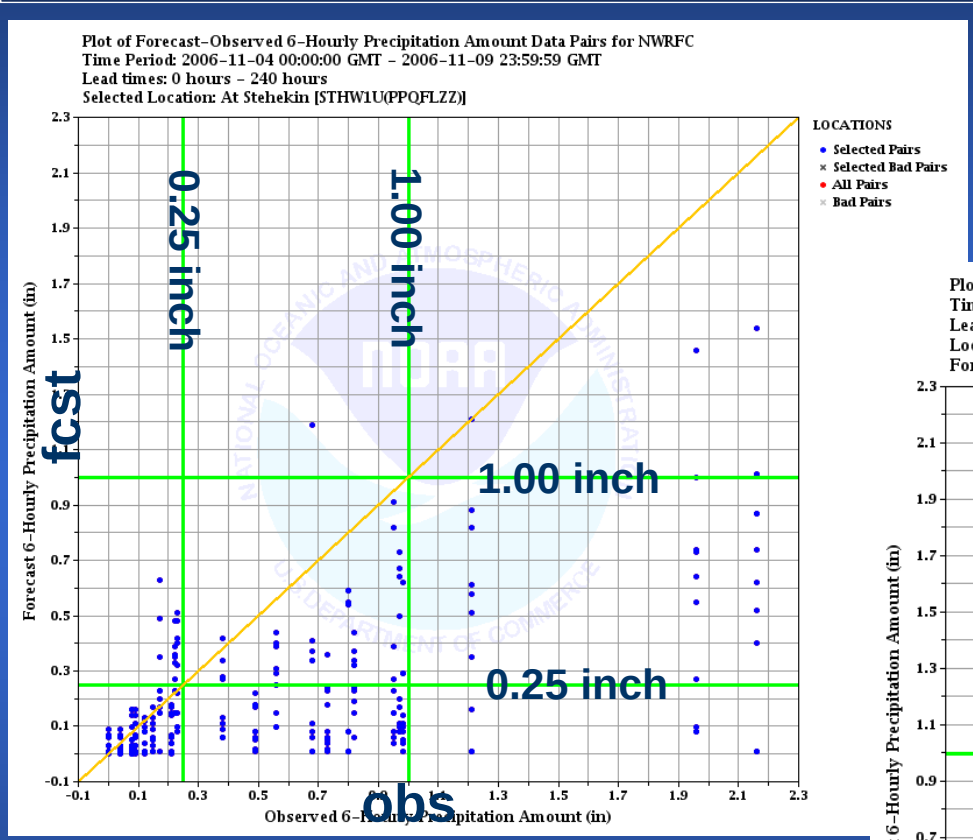
Stehekin – Time Series and Scatter Plots for Observed and Forecast Stage

Analysis Period: Nov 4-9, 2006

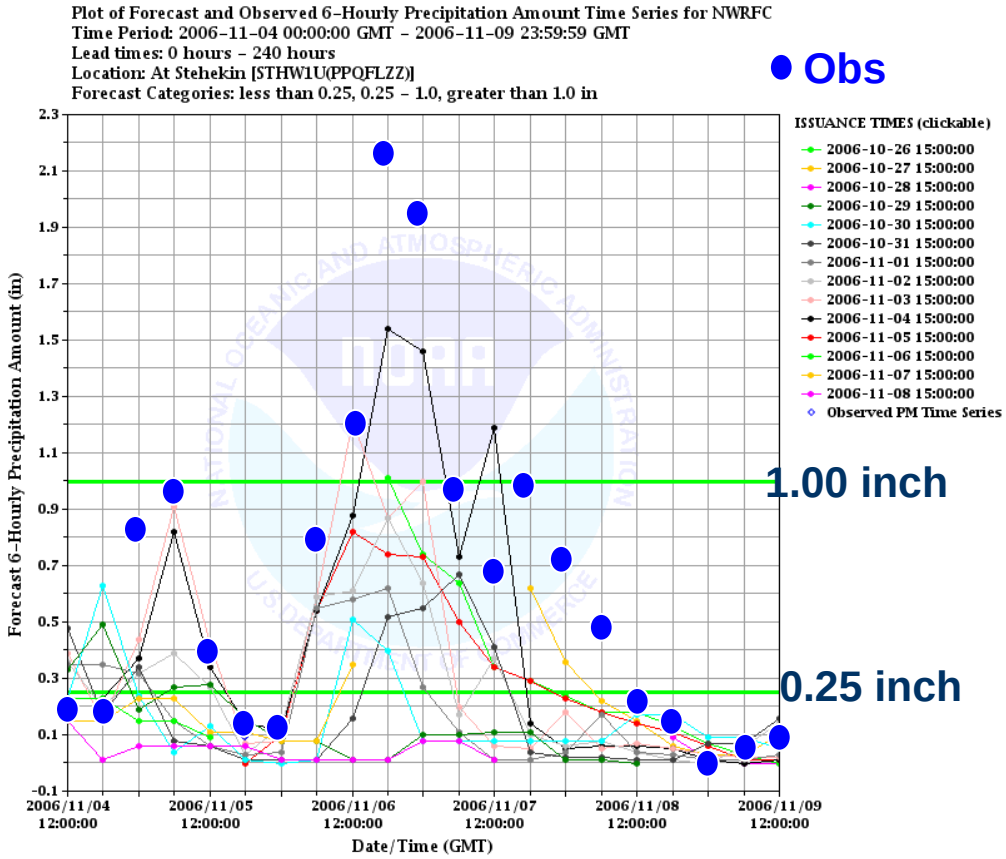




Upper zone MAP/FMAP (6hr values) – Categorical Scatter and Time Series Plot.

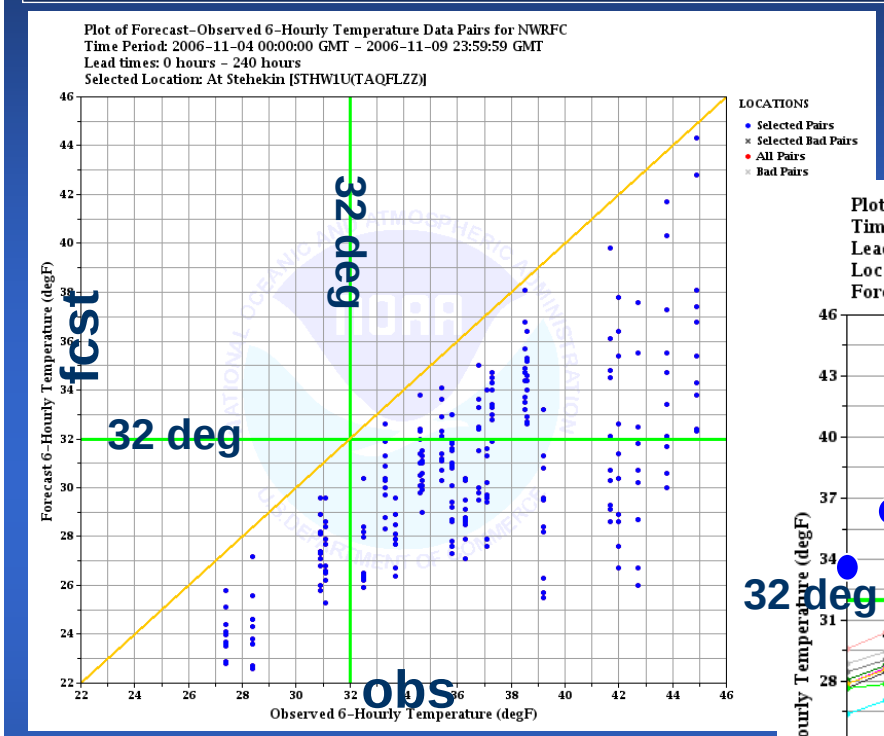


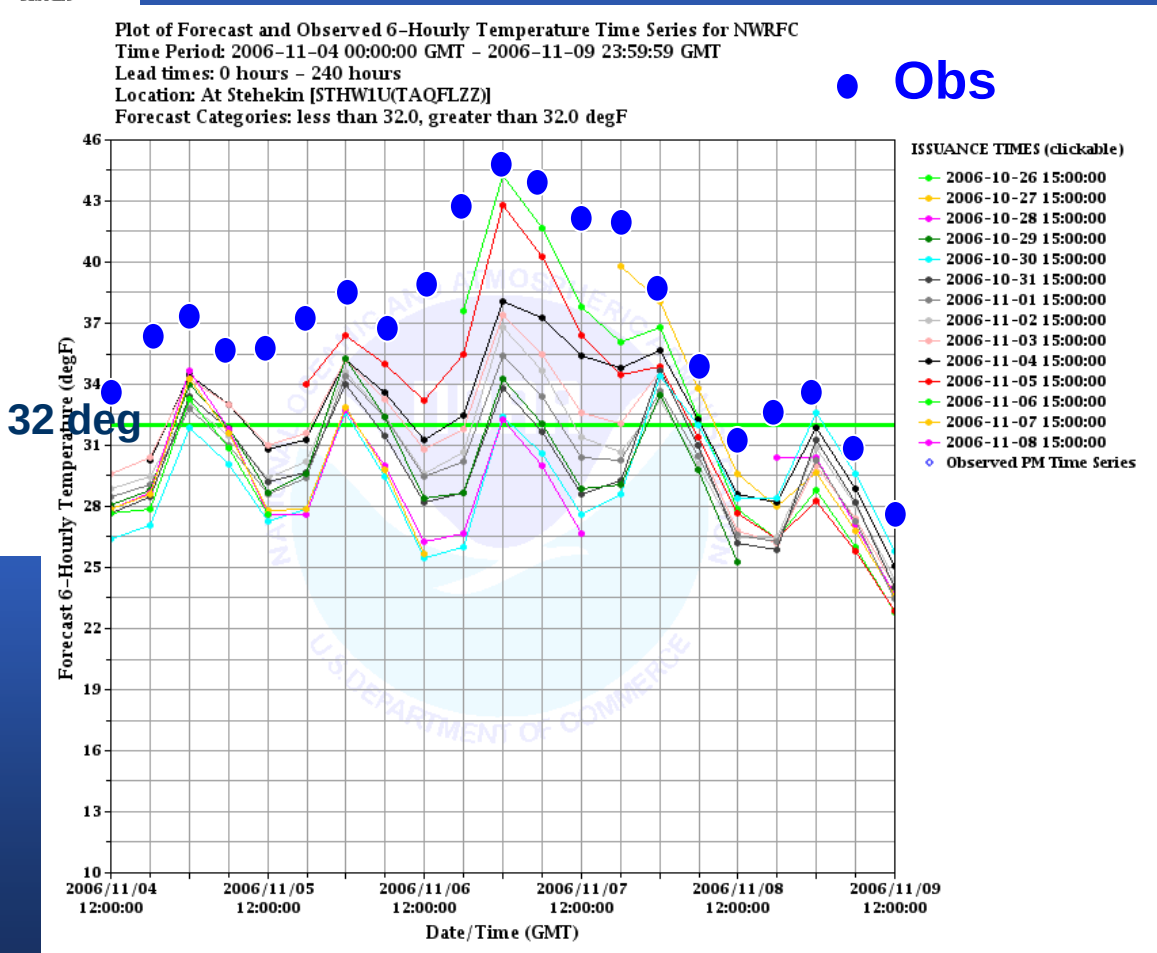
Analysis Period: Nov 4-9, 2006



- Forecasts for values which came in less than 0.25" were fairly "evenly scattered"
- Consistent Low Bias for values > 0.25"
- Note: Lower zone showed similar pattern.

Upper Zone MAT/FMAT (6 hour)– Scatter and Time Series Plots



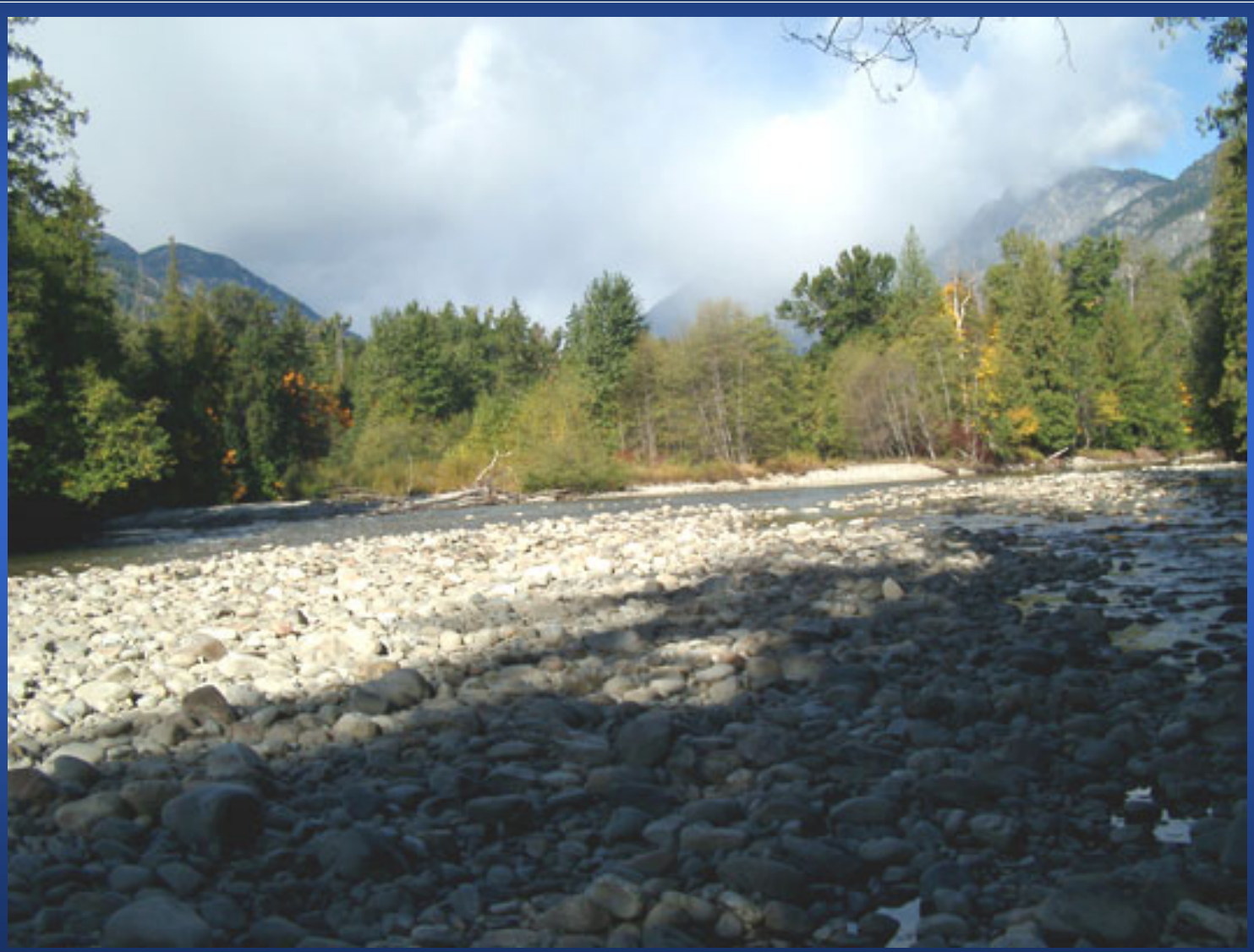


- Consistent Low Bias for all values
- Note: Lower zone showed similar pattern.

IVP Analysis Summary

- Stage and Flow forecast were generally low throughout the period. Many issuances near the crest produced high biased recession forecasts.
- FMAP was very low compared to MAP throughout the period. There were only scattered exceptions.
- FMATs were very bias on the “low side” as well.
- Persistence forecasts were superior to FMATs at times for the short lead times and the very long lead times.
- Forecasts “generally” improved with shorter lead times.
- Statistics/plots are extremely sensitive to the analysis period, lead-time, and number of samples.
- Relative error contribution of FMAP, FMAT, antecedent conditions, modeling, forecaster, etc. cannot really be assessed with IVP
- There is some application for single event analysis in IVP. However, it is probably better suited for multi-event analysis in general.

IFP “RE-ANALYSIS”:



NOAA NWS ~ NORTHWEST RIVER FORECAST CENTER

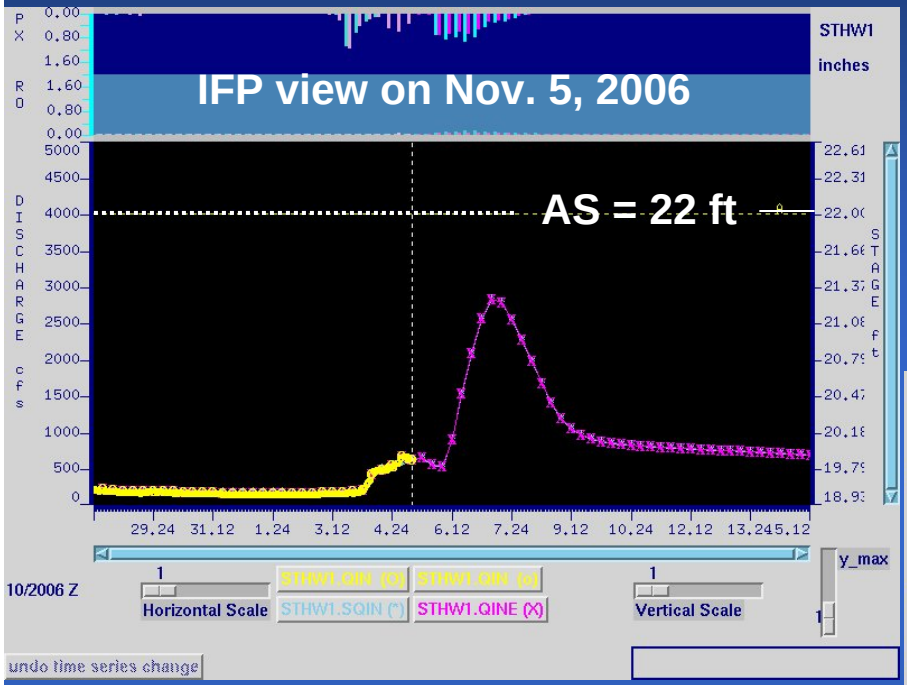
STHW1 – Verification Analysis using IFP with Archived OFS Files: Goals/Approach

- Look at the event in a forecast mode (2 days prior to the crest).
  - Analyze soil moisture and snow model states
  - Look to see if any forecaster mods could have negatively/positively impacted the forecast
- Look at the event in an "observed" mode (i.e. 2 days after the peak)
  - Look at a "no mods" scenario
  - Try to isolate biggest error contributor (soil moisture, snow, FMAT, FMAP, etc.)
  - Determine the minimum and "most likely" mods required to simulate the storm in an "observed" mode.

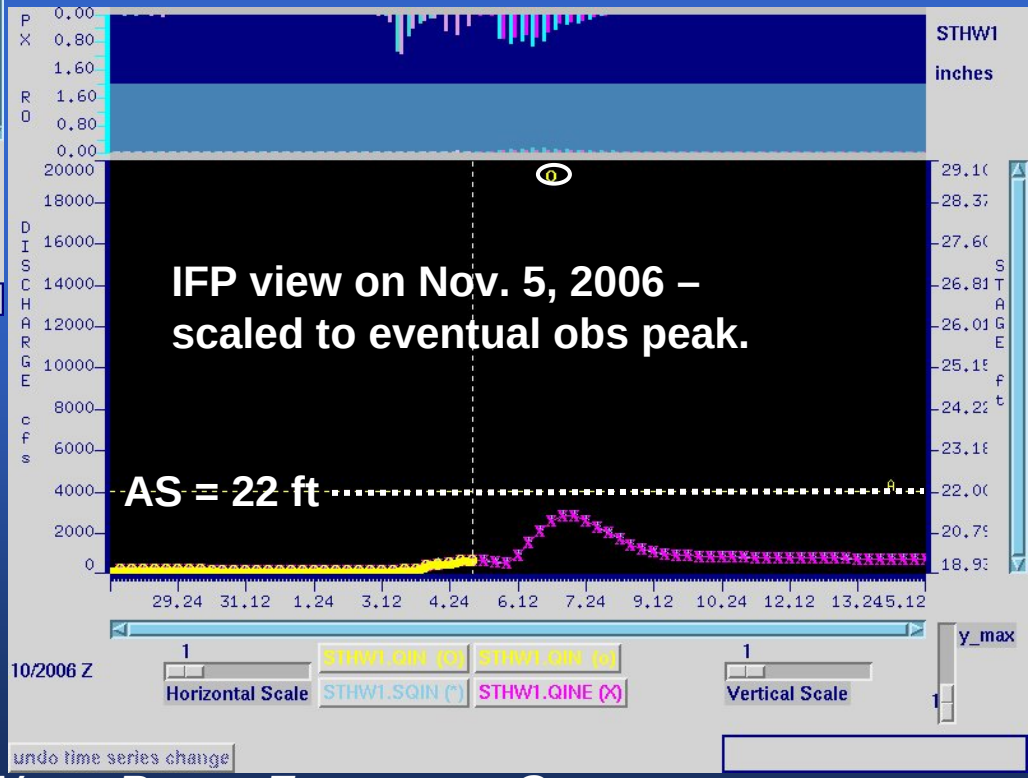
Antecedent Conditions : Were the initial Snow and Soil Moisture conditions reasonable?

- **The initial SWE on Nov 1 was:**
  - **Upper zone: 2.0” [mean elev = 6160’]**
  - **Lower zone: 0.1” [mean elev = 3800’]**
  - **SNOTEL sites in the area had very similar SWE readings on Nov 1.**
- **Simulated snow “built” to a max of around 6.5” in the upper zone on Nov 7th (once again…this was representative based on area SNOTEL sites).**
- **Simulated initial simulated soil moisture states were reasonable. Observed flow and SAC-SMA state climatology data were used as references for this assessment.**

IFP Forecast View on Nov. 5, 2006

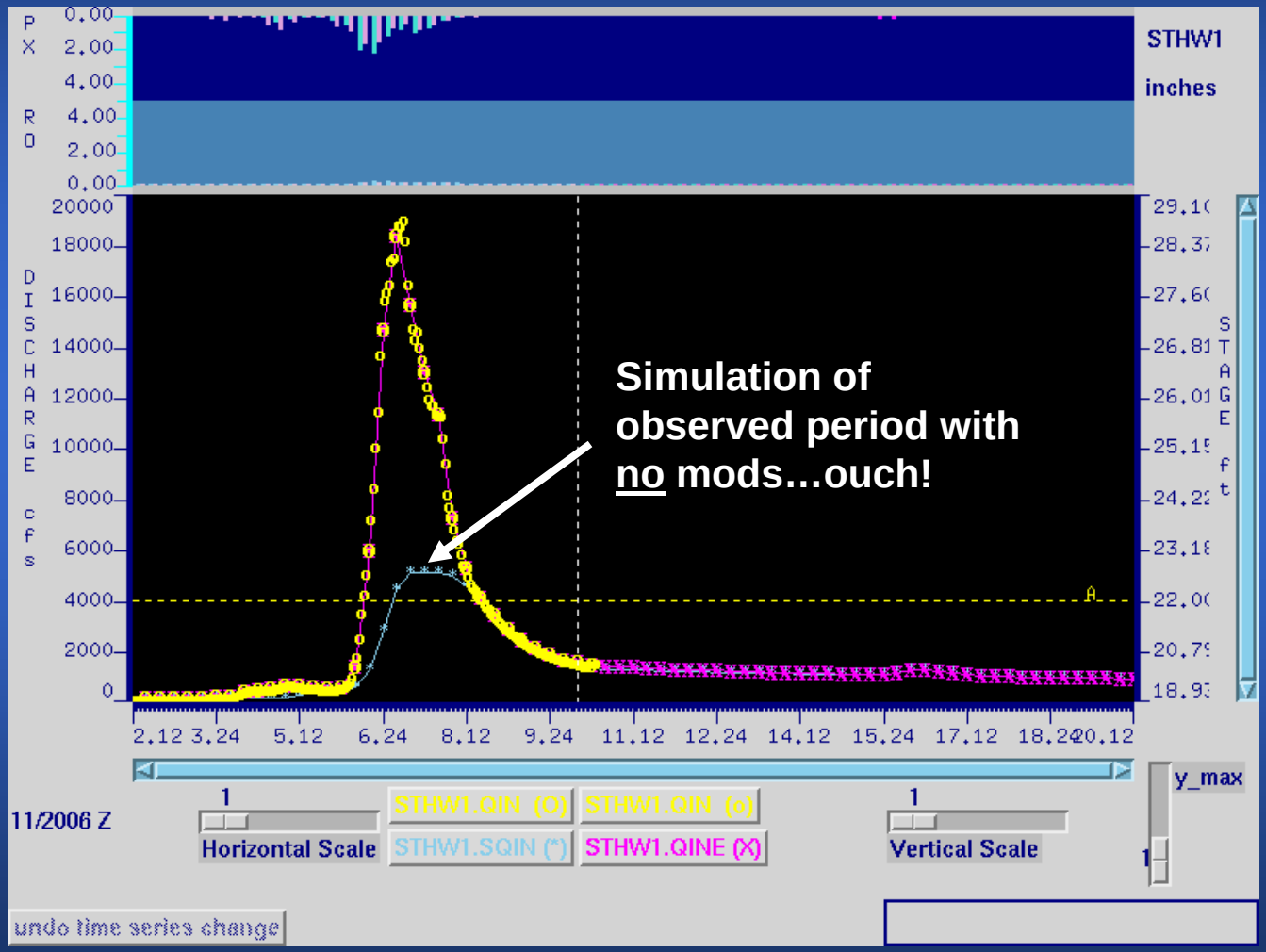


- Model states appear to be adjusted properly.
- There were some forecaster mods which would impact the precipitation typing (i.e. more rain, less snow) during the event (Good move!).



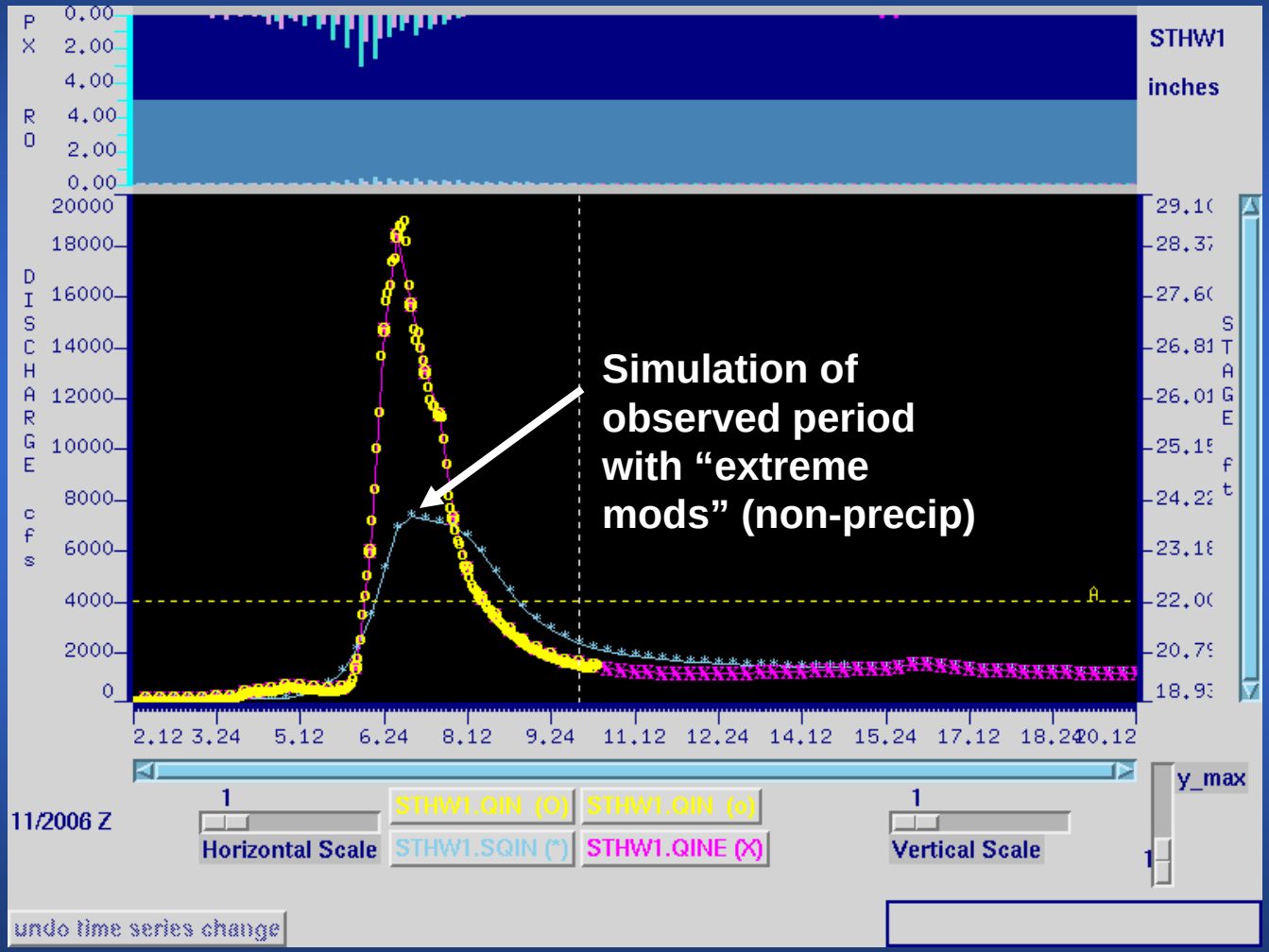
- Quite frankly….the forecast looks reasonable given the information available at the time.

Post event analysis using IFP: Nov 10th OFS Files – No Mods



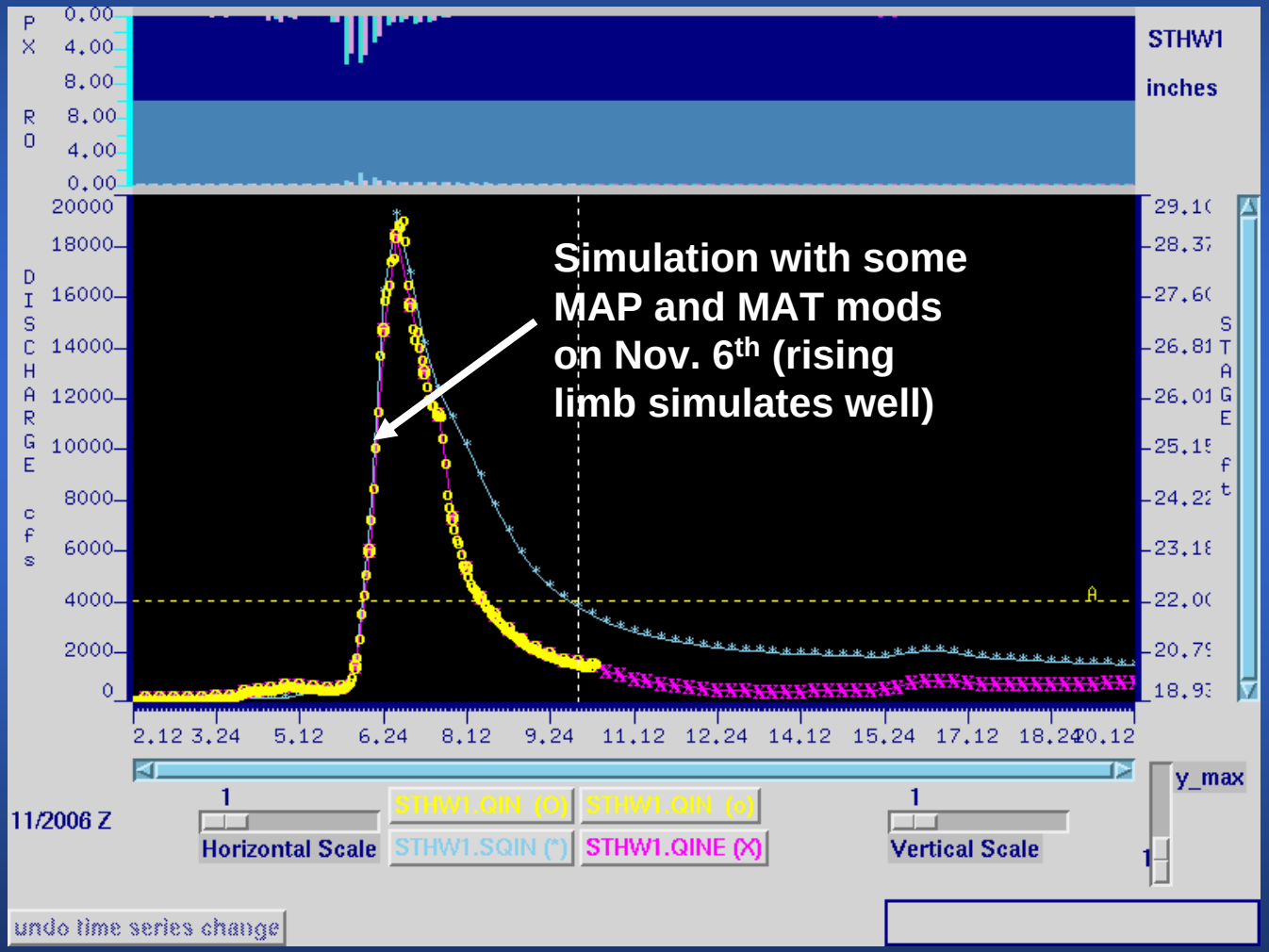
Combination of "observed" MAT and MAP was entirely inadequate!

Results of Post Event analysis - “Extreme Mods” (non-precip)



- Nov 5 : Filled Tension Water “Buckets”
- Nov 5-8 : Added 20 degrees to all MATs
- Nov 5-8 : Increased UADJ to maximum (wind effects for rain on snow)
- Result: Simulation is still way “underdone”!
- Probable cause: MAP too low! (observed point precipitation is either bad or not producing representative MAPs)

Results of Post Event analysis – What would it take to match this "rise"?

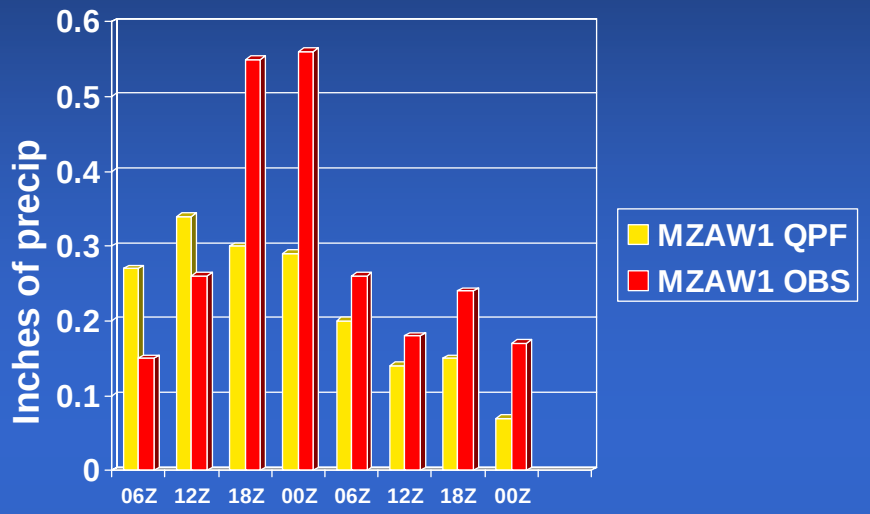


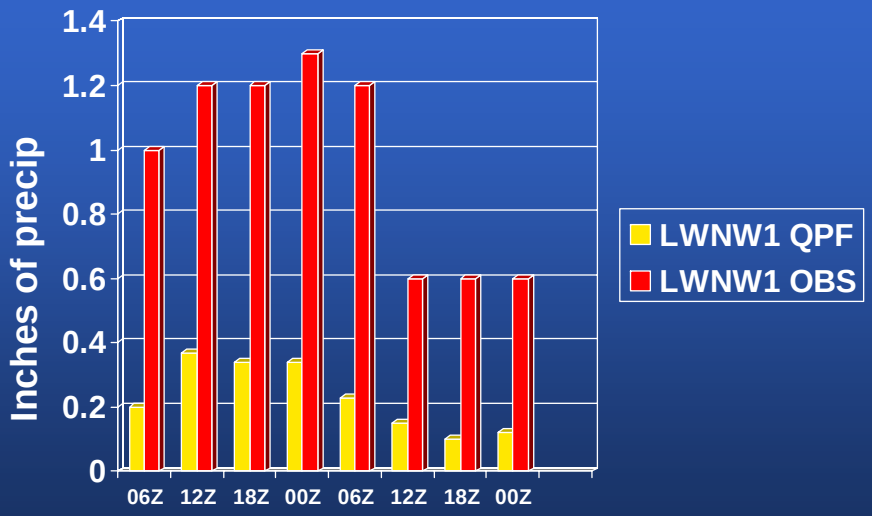
- Re-distribute heavy precip (13 inches) to concentrate in an 18 hr period starting on Nov 6 at 12Z
- Add 10 degrees to all MATs from Nov 6 06Z-12Z
- Result: rise can be simulated. Recession would have problems without additional mods to reduce subsequent precipitation.

IFP Summary for STHW1

- Using a very “quick and crude” sensitivity analysis, it is very obvious that forecast mean areal precipitation (FMAP) produced the largest errors in the forecast.
- Low-biased mean areal temperature (FMAT) forecasts had an impact as well. However, the impact was minor compared to the FMAP.
- Antecedent conditions (model states) were reasonable.
  - Simulated snow water equivalent (SWE) values compared favorably with SNOTEL “ground truth”.
  - SAC-SMA state values were reasonable based on Fall precip, observed flow, and Climatological averages.
- “Observed” MAP and MAT values needed significant modification in the model to simulate the observed flow time series. Point network was not representative.
- Although “playing” with archived OFS files using IFP is user friendly and can yield some interesting results, the impacts of point precipitation and temperature data cannot be assessed via this method.
- “Future” temperature Mods produced by a forecaster had a positive impact on the forecast. No other mods had any significant impact on the forecast in a positive or negative way.

QPF points used to produce STHW1 FMAPs – 6 hour totals (inches): 11/06 06Z – 11/08 00Z



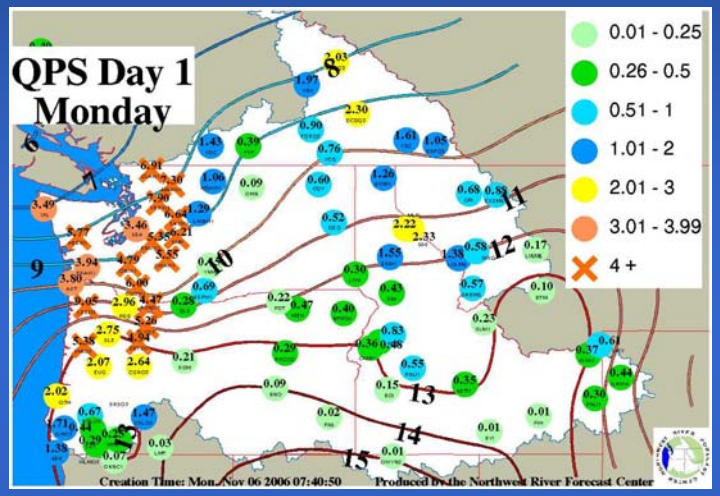


- Prior to 2007:
  - The QPF network was a manageable sub-set of the Calibration precipitation network. Each FMAP was computed as a weighted sum of these sites (additional normalization was performed relating annual normals). This sparse network was known to produce "unrepresentative" FMAPs at times.
- After 2007:
  - Currently, the GFE tool is used to produce QPF and the network has expanded to include all calibration sites (i.e. the same stations and absolute weights are used for MAP and FMAP).
- It's worth noting:
  - Since this expansion of the forecast network, neither of these 2006 QPF points are used to generate mean areal time series for the Stehekin!

Contrasting QPF Point Networks

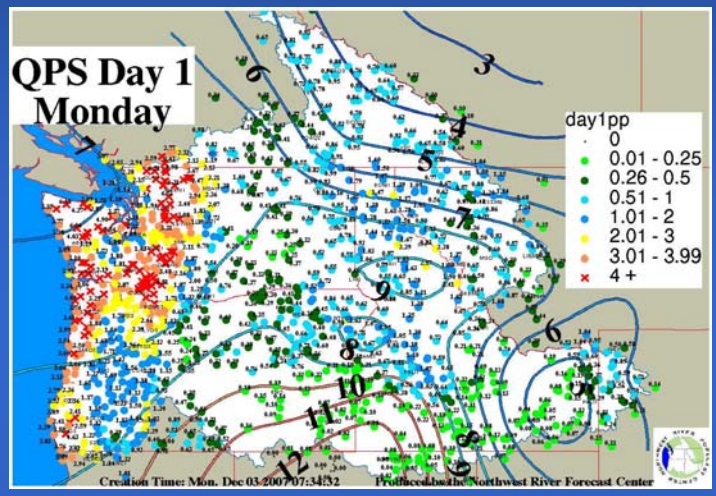
QPF Plot generated on Monday Nov 6, 2006 – "old" network.

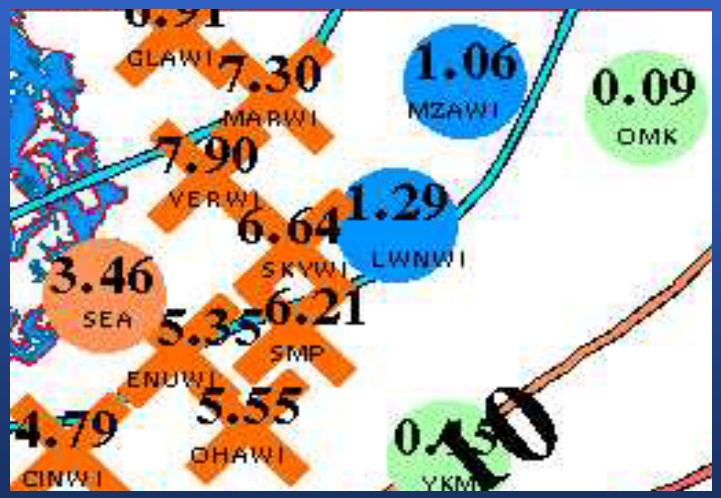
QPF Plot generated on Monday Dec 3, 2007 – calibration station network.



← OLD

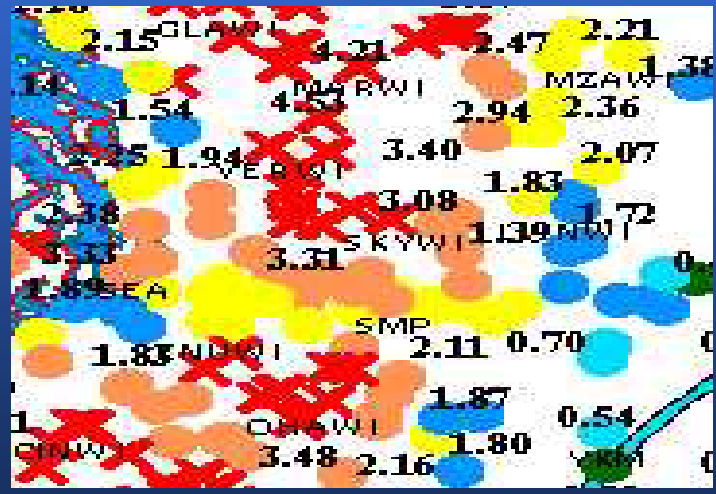
NEW →





← OLD

NEW →



NOAA NWS ~ NORTHWEST RIVER FORECAST CENTER

Conclusions

- IVP and IFP (re-analysis) have definite strengths and weaknesses.
- IVP is a very powerful tool that appears to have more utility in a multi-event analysis rather than an individual storm study.
- Time series and scatter plots are (by far) the most useful options within IVP for an individual storm analysis.
- IFP is very useful in a post event analysis (requires routine archive of OFS files).
- Using a combination of IVP and IFP can be an effective approach to performing Event Verification. With this combo (or something similar), we can possibly point out inherent problems with our forecasts that could be fixed (i.e. Possibly Improve Forecasts!)

Recommendations

- Exercising the IVP has opened our eyes to many possibilities for verification analysis. This includes current capabilities and possible future enhancements. (Just like all software, I'm sure that the list is growing).
- Recent improvements to the Archive system (size, speed) are welcome. However, IVP still encounters performance issues during the analysis of large datasets. Unfortunately, this is where IVP analysis is most relevant. Workarounds have helped some, but the underlying problem needs to be identified and addressed.
- Archived data is critical for many RFC programs (ops, verification, calibration, ensembles, climatologies, etc.). The Archive must be able to communicate with CHPS/FEWS.
- Effective Data viewing and quality control tools for the Archive are very important "pieces" of a Verification system. We need to continue to improve those tools.
- Migration to the new forecast system (CHPS/FEWS) will require a replacement for the IFP/OFS post-event analysis method. Current capabilities within FEWS could possibly satisfy this need. This possibility needs to be evaluated.

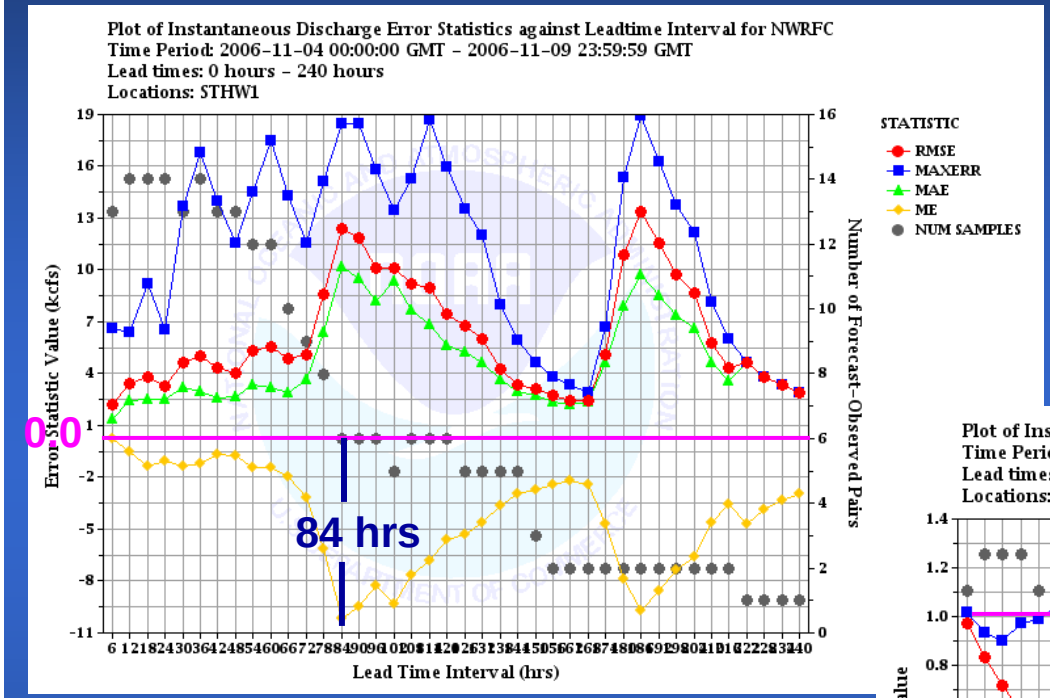
Stehekin River Valley

NOAA NWS ~ NORTHWEST RIVER FORECAST CENTER

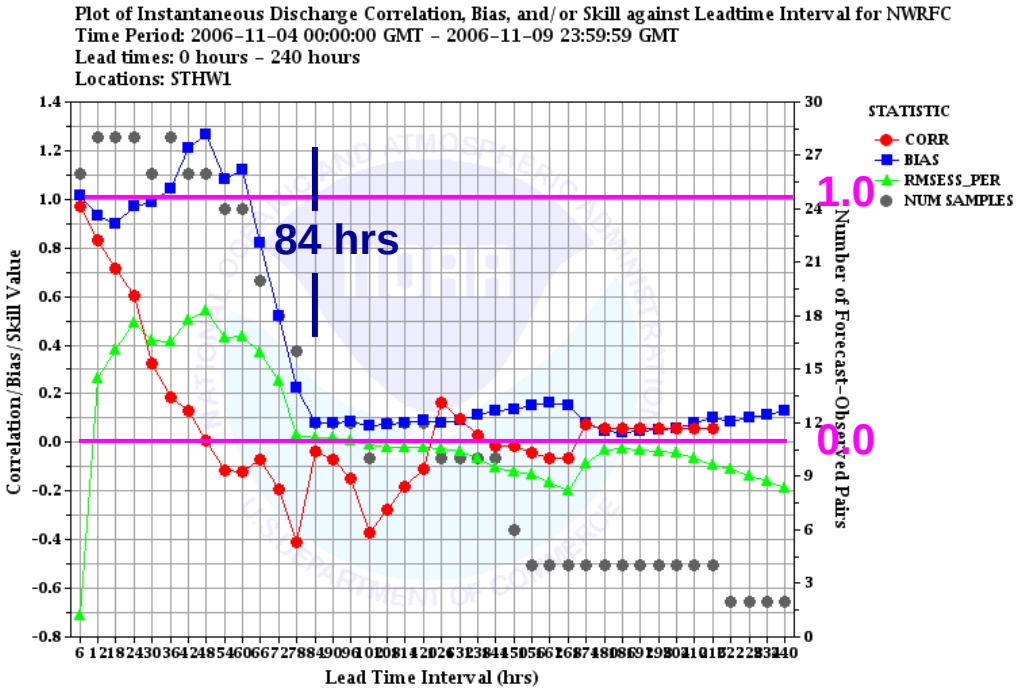
Possible Future Work

- **Unanswered Questions:**
  - The mean areal data (forecast and observed) can be “unrepresentative” even if the point data from which they are generated have verified perfectly. Why weren’t the point date representative in this case? Did high winds lead to gage catch deficiencies?
  - At the time of this event, we were using a limited point network to produce forecast precipitation. Our current QPF/Temperature forecasting method employs GFE tool and produces a much more “dense” network of point forecasts. Are current and future forecasts likely to improve as a result of this change?
  - Could quality controlled obs temp and precip estimates be flawed in this basin?
  - Are these event forecast biases “typical” in this basin?
  - Were these forecast problems widespread or isolated to the Stehekin during this event?
- Potential Future Work:
  - Look at other storms for this basin to determine if we have any forecast trends or identifiable biases present in all cases.
  - Perform a similar verification analysis on other “nearby” river basins.
  - Set up point verification in IVP and perform routine analysis (temp & precipitation)

Stehekin Flow Forecast Statistics

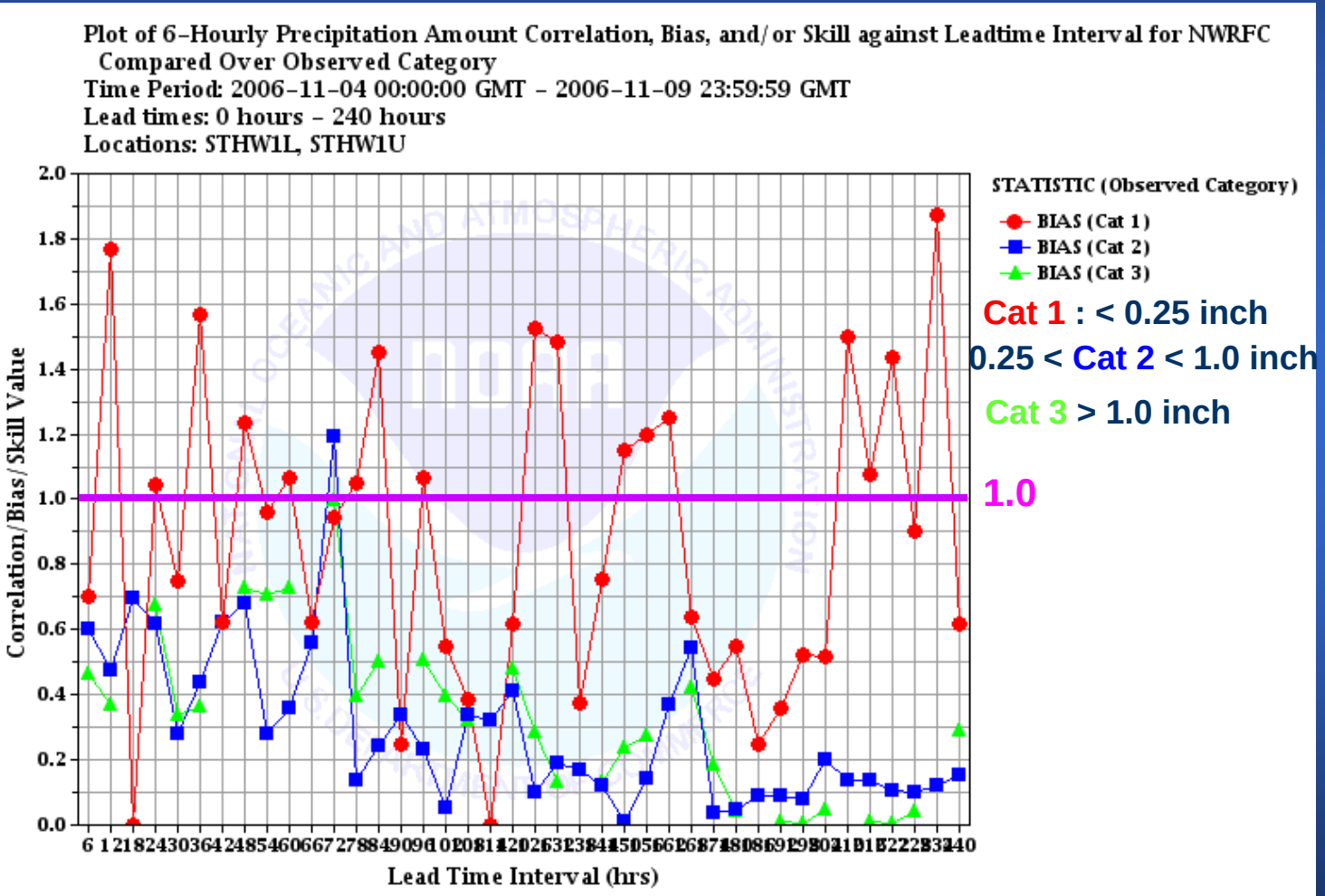


- Consistent Low Bias on Rising Limb with the exception of last few forecasts
- Recession forecasts had high Bias

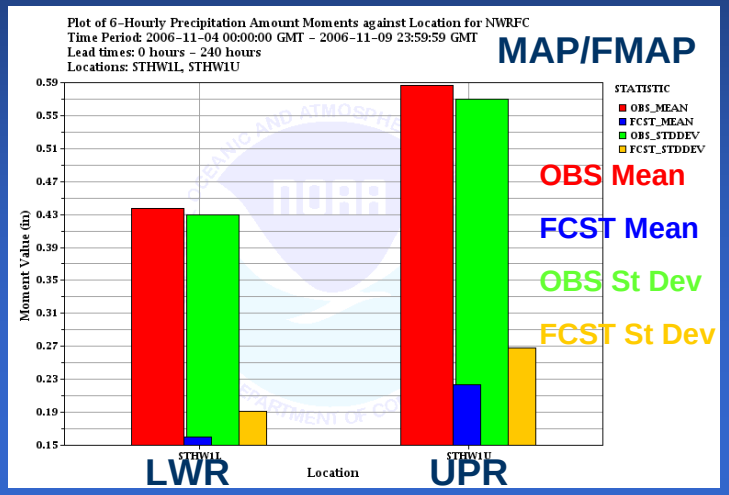


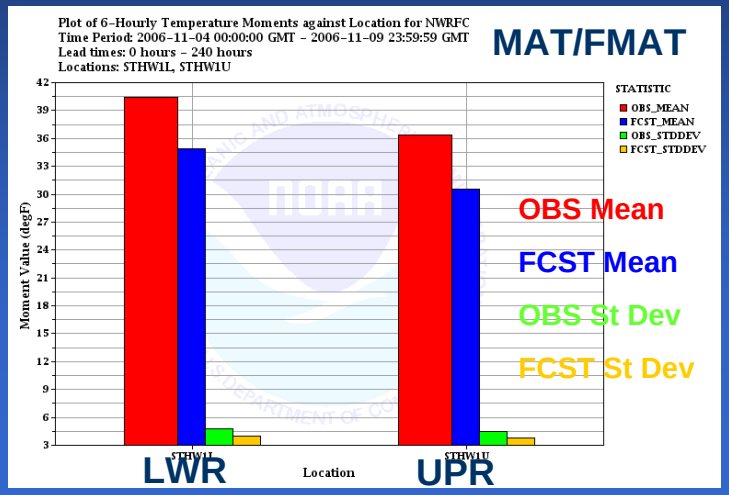
- ME was negative throughout most of the analysis period.
- Forecasts showed "general" statistical improvement starting at 84 hour lead-time

Categorical Bias for MAP (U & L)



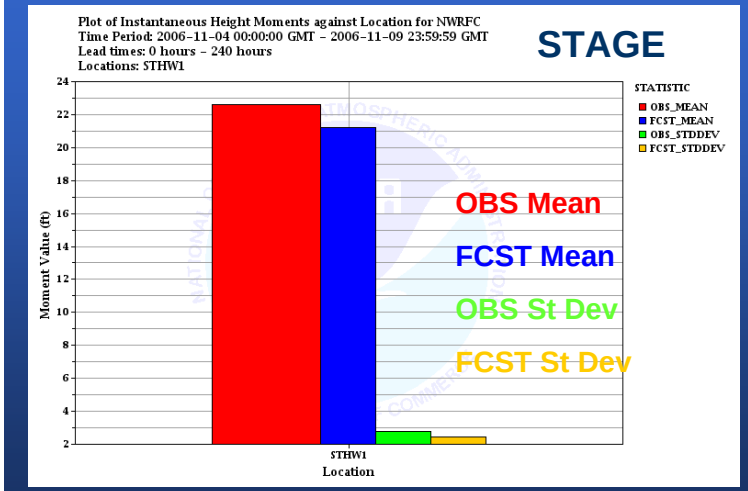
Moments for all Forecasts during event

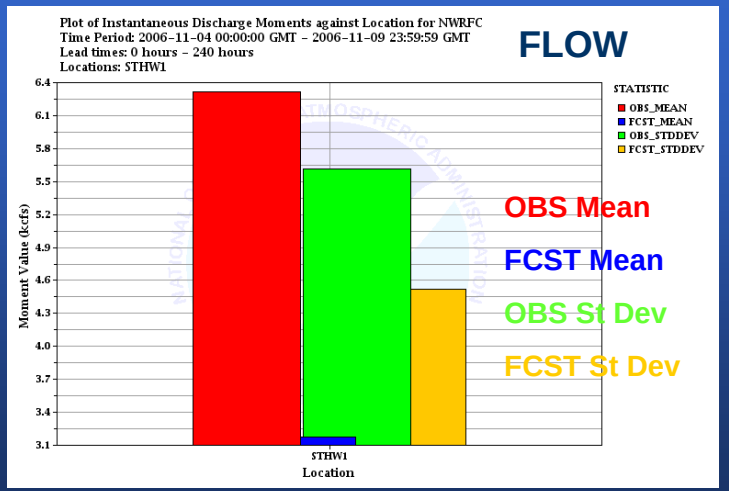




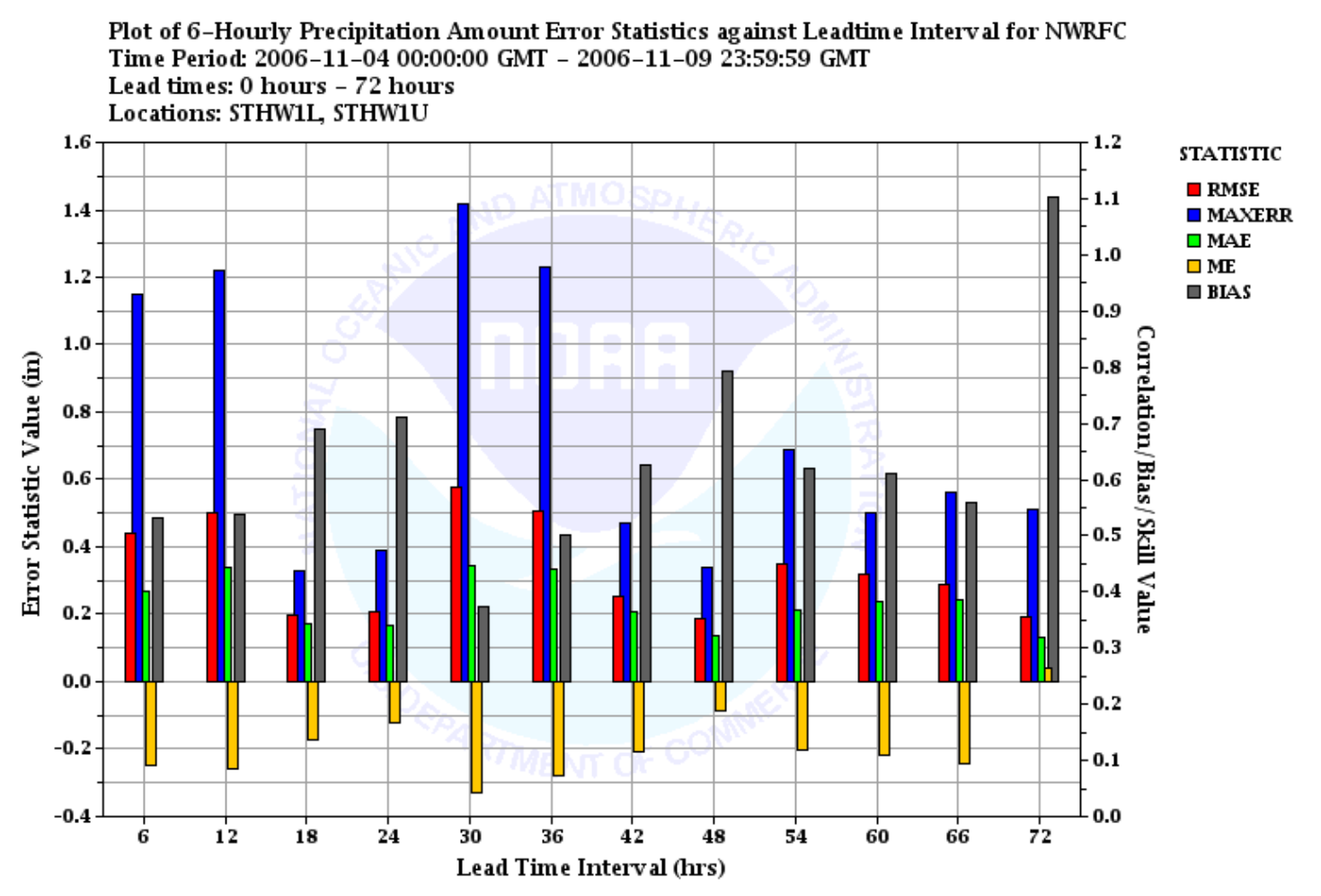
Bottom Line:

Everything was low.

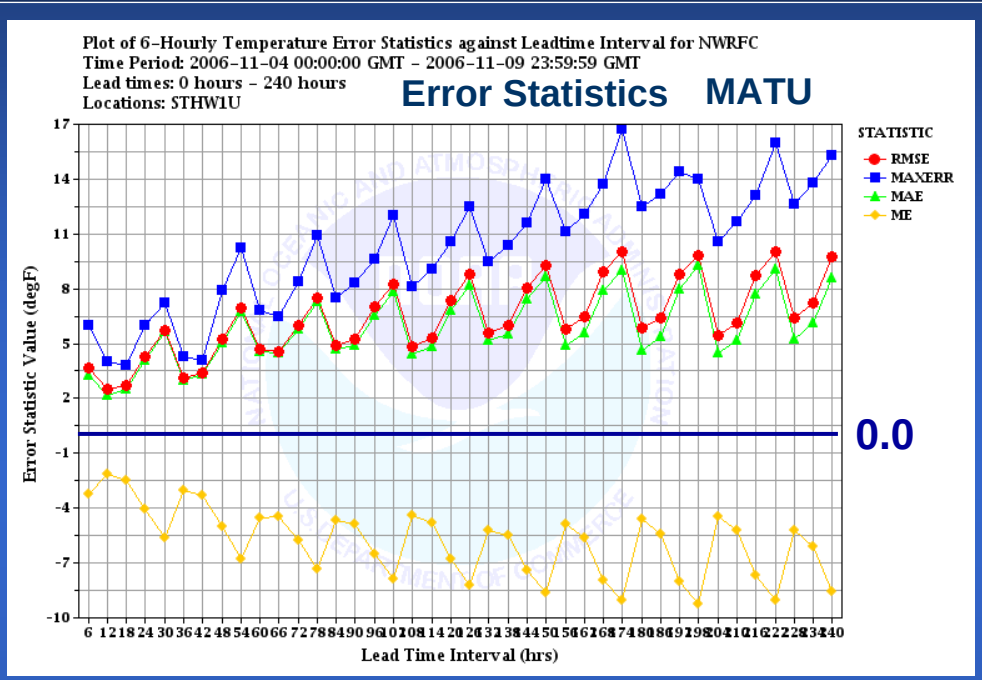


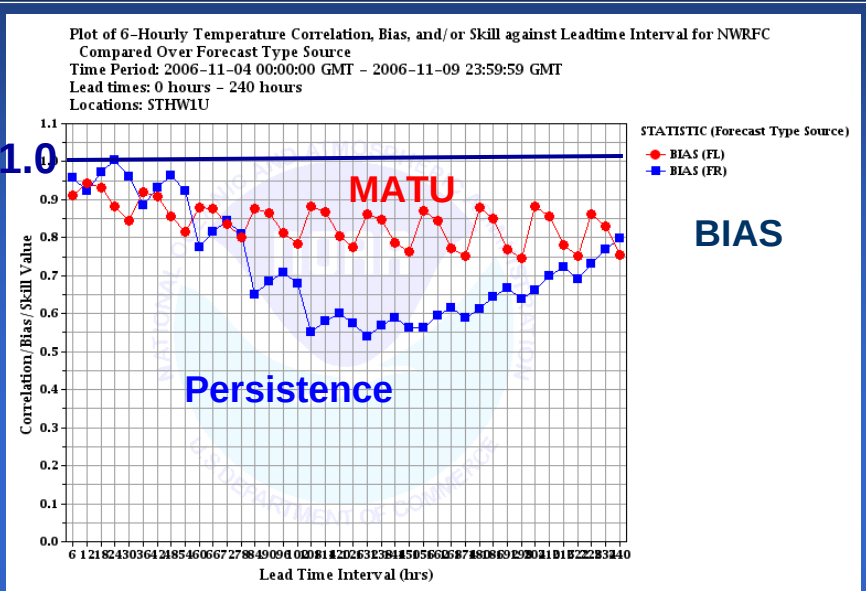


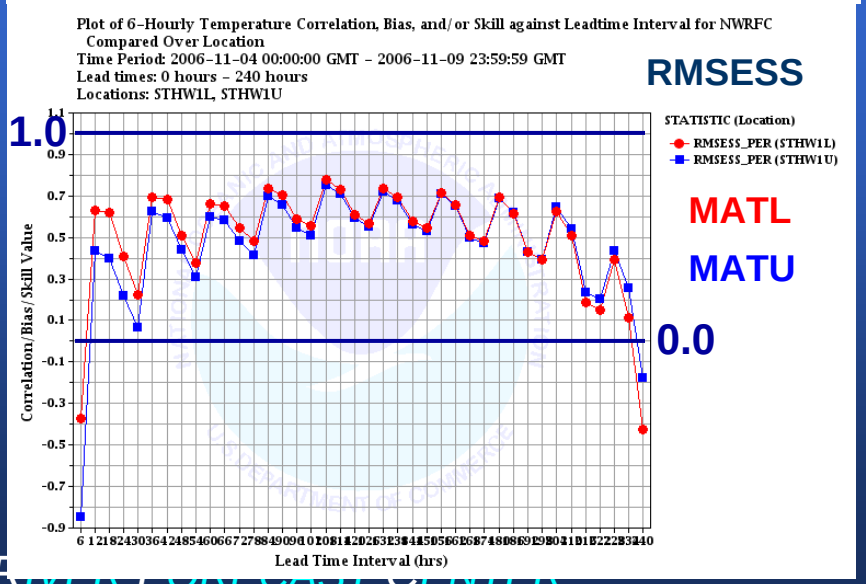
MAP Upr & Lwr Error Stats and Bias : 3 Day Lead Time



MAT Analysis







Time Series Plot for:
Observed and Forecast Stage (STHW1)