Western Water Verification

www.nwrfc.noaa.gov/westernwater/beta

Lisa Holts Kevin Werner

Outline

- Verification capabilities
- Graphing capabilities
- Example

Motivations

- Improve understanding of uncertainty from forecast tools
 - How good are our forecasts?
 - How much value do human forecasters add?
 - Is one tool better than another?
 - How do the above questions depend on lead time? Amount of runoff?
- Can we use what we learn to:
 - Improve forecasts?
 - Make forecast process more efficient?
 - Improve collaboration?
 - Convey uncertainty to our users?

NOAA/NWS Water Resource Forecasts	water supply map	water supply forecasts	ensemble forecast	verification	climate scenarios	data checkout	about western water
Location: Invalid Location Search: Browse: Select State	Go Select State	First 💌 Select	State First 💌			change location	clear location
	Locati	on must be selecte	ed before verificati	on data can be dis;	played		
Water Supply Forecast Application - v	ersion 2.0 (beta 1)					Natior	al Weather Service

- Search By
 - Station ID
 - State, River, Location
- Location
 - Change
 - Clear
 - Return to main menu

NOAA/NWS Water Resource Forecasts	water supply map	water supply forecasts	ensemble forecast	verification	climate scenarios	data checkout	about western water
Location: Blue at Dillon Res, Col	lorado (DIRC2 - CBR	FC)			1	change location	clear location
Seasonal Forecast	Verification a	nd Analysis					
	1 Select	t Period:	Select Pe	eriod 💌			
	2 Selec	t Years:	Select Pe	eriod First 💌			
	3 Select	<u>t Data Sources:</u>	 ✓ Coordi ✓ Nation ✓ Natura ✓ Statisti ✓ Califor ✓ Ensen ⊂ Ensen 	nated Forecast al Weather Service I Resource Conser cal Water Supply nia Department of V nble Streamflow Pro- nble Streamflow Pro-	wation Service Water Resources ediction - Empirica ediction - Normal ediction - Lognorn ediction - Wakeby ediction - Logwieb ediction - Weibull ediction - Loglogis	al nal bull	



- Historical
- O Streamflow Histogram
- C Scatterplot
- O Mean Absolute Error (Lead Time)
- Mean Absolute Error (Years)
- C Root Mean Sqaured Error (Lead Time)
- C Root Mean Sqaured Error (Year)
- C Skill Score (Lead Time)
- O Skill Score (Years)
- O Percent Difference
- Probability of Detection Above Threshold
- O Probability of Detection Below Threshold
- C False Alarm Rate Above Threshold
- C False Alarm Rate Below Threshold
- C Rank Histogram
- C Rank Histogram (Lead Time)
- C Climate Variability
- C Contingency Table

Load Statistics

About Forecast Verification

Forecast verification provides meaningful information about the quality of forecasts. Verification is an important for assessing past forecast performance and providing information about current forecast confidence. The seasonal forecast verification application allows users to obtain relatively simple plots and statistics explaining past performance. The application allows the user to choose (1) a forecast period (season), (2) the forecast years, (3) forecast types, and (4) verification metric. All data supporting the plots are archived in a database and may be accessed separately through the data checkout application as well. Site Options Previous 5 Locations Viewed GBYC2 Print Graph Display Raw Data

NOAA/NWS Water Resource Forecasts	water supply map	water supply forecasts	ensemble forecast	verification	climate scenarios	data checkout	about western water
Location: Blue at Dillon Res, Co	lorado (DIRC2 - CBR	FC)				change location	clear location
Seasonal Forecast	Verification a	nd Analysis					
	1 Select	t Period:	Select Pe		•Sele	ect Forec	ast Period
	2 Select Years:			Select Period First			vailable
	3 <u>Select</u>	t Data Sources:	☑ Coordi ☑ Nation	nated Forecast al Weather Service	•Clim	atology /	Available
			✓ Natura ✓ Statisti	l Resource Conser cal Water Supply	rvation Service		
			Califor	nia Department of \ hble Streamflow Pro	Water Resources ediction		
			Ensen	nble Streamflow Pro hble Streamflow Pro	ediction - Empirica ediction - Normal ediction - Lognorm	nal	
			Ensen	nble Streamflow Pre table Streamflow Pre	ediction - Lognom ediction - Wakeby ediction - Logwieb	ull	
			Ensen	nble Streamflow Pre nble Streamflow Pre	ediction - Weibull ediction - Loglogis	tic	





	A Select Plot Type:	Historical
	U	O Streamflow Histogram
		C Scatterplot
	Salact Plat Type	Mean Absolute Error (Lead Time)
	•Select Flot Type	Mean Absolute Error (Years)
		C Root Mean Sqaured Error (Lead Time)
	 Historical By Default 	C Root Mean Sqaured Error (Year)
		Skill Score (Lead Time)
		Skill Score (Years)
		O Percent Difference
		O Probability of Detection Above Threshold
		O Probability of Detection Below Threshold
\mathbf{X}		C False Alarm Rate Above Threshold
		C False Alarm Rate Below Threshold
		C Rank Histogram
		C Rank Histogram (Lead Time)
		Climate Variability
		C Contingency Table
		Last Of the
		LCCC STATISTICS

About Forecast Verification

Forecast verification provides meaningful information about the quality of forecasts. Verification is an important for assessing past forecast performance and providing information about current forecast confidence. The seasonal forecast verification application allows users to obtain relatively simple plots and statistics explaining past performance. The application allows the user to choose (1) a forecast period (season), (2) the forecast years, (3) forecast types, and (4) verification metric. All data supporting the plots are archived in a database and may be accessed separately through the data checkout application as well.

Site Options Previous 5 Locations Viewed <u>GBYC2</u>

Print Graph Display Raw Data



- O Historical
- C Streamflow Histogram
- O Scatterplot
- C Mean Absolute Error (Lead Time)
- O Mean Absolute Error (Years)
- C Root Mean Sqaured Error (Lead Time)
- C Root Mean Sqaured Error (Year)
- C Skill Score (Lead Time)
- O Skill Score (Years)
- O Percent Difference
- C Probability of Detection Above Threshold
- O Probability of Detection Below Threshold
- C False Alarm Rate Above Threshold
- C False Alarm Rate Below Threshold
- C Rank Histogram
- C Rank Histogram (Lead Time)
- C Climate Variability

O Contingency Table

Load Statistics

About Forecast Verification

Forecast verification provides meaningful information about the quality of forecasts. Verification is an important for assessing past forecast performance and providing information about current forecast confidence. The seasonal forecast verification application allows users to obtain relatively simple plots and statistics explaining past performance. The application allows the user to choose (1) a forecast period (season), (2) the forecast years, (3) forecast types, and (4) verification metric. All data supporting the plots are archived in a database and may be accessed separately through the data checkout application as well.

Site Options

Previous 5 Locations Viewed <u>GBYC2</u>

Load Statistics

Print Graph Display Raw Data



- Historical
- O Streamflow Histogram
- O Scatterplot
- C Mean Absolute Error (Lead Time)
- O Mean Absolute Error (Years)
- C Root Mean Sqaured Error (Year)
- C Skill Score (Lead Time)
- C Skill Score (Years)
- O Percent Difference
- C Probability of Detection Above Threshold
- C Probability of Detection Below Threshold
- C False Alarm Rate Above Threshold
- C False Alarm Rate Below Threshold
- C Rank Histogram
- C Rank Histogram (Lead Time)
- C Climate Variability
- C Contingency Table

Load Statistics

About Forecast Verification

Forecast verification provides meaningful information about the guality of forecasts. Verification is an important for assessing past forecast performance and providing information about current forecast confidence. The seasonal forecast verification application allows users to obtain relatively simple plots and statistics explaining past performance. The application allows the user to choose (1) a forecast period (season), (2) the forecast years, (3) forecast types, and (4) verification metric. All data supporting the plots are archived in a database and may be accessed separately through the data checkout application as well.

Site Options

Previous 5 Locations Viewed GBYC2

Print Graph Display Raw Data

○ Root Mean Sqaured Error (Lead Time) ● Changes based on Statistic to explain graph

About Verification

•General description



- Historical
- O Streamflow Histogram
- O Scatterplot
- C Mean Absolute Error (Lead Time)
- O Mean Absolute Error (Years)
- C Root Mean Sqaured Error (Lead Time)
- C Root Mean Sqaured Error (Year)
- C Skill Score (Lead Time)
- C Skill Score (Years)
- O Percent Difference
- C Probability of Detection Above Threshold
- C Probability of Detection Below Threshold
- C False Alarm Rate Above Threshold
- C False Alarm Rate Below Threshold
- C Rank Histogram
- C Rank Histogram (Lead Time)
- C Climate Variability
- C Contingency Table

Load Statistics

About Forecast Verification

Forecast verification provides meaningful information about the quality of forecasts. Verification is an important for assessing past forecast performance and providing information about current forecast confidence. The seasonal forecast verification application allows users to obtain relatively simple plots and statistics explaining past performance. The application allows the user to choose (1) a forecast period (season), (2) the forecast years, (3) forecast types, and (4) verification metric. All data supporting the plots are archived in a database and may be accessed separately through the data checkout application as well.



Previous Sites

•Remembers last 5 visited

Go to Data Checkout

•Print Graph



water supply

forecasts

water supply



- Statistic
- Forecast
 Type/Data
 Source
- Time Scale
- Threshold
- Data Source
 Information
- Graph Options



forecast

verification

The historical plot displays the forecasted streamflow for the selected agencies by month and year for the forecast period. The forecast period is the month range that a forecast is valid for and the months that the observed streamflow is summed over. The observed streamflow period of record maximum and minimum streamflow is displayed as well as the threshold value (historical average is default).

Water Supply Forecast Application - version 2.0 (beta 1)

NOAA/NWS Water

Resource Forecasts

National Weather Service

Display Raw Data

Print Graph

western water

Options

Statistic

• Statistic

- 18 Statistic choices
- Graph change on click
- Mouse-over displays info about graph

• Group Titles

 Mouse and click to display other options

The Historical displays forecast streamflows for the selected period from all agencies and observed streamflows from past years while referencing the historical maximum, minimum, and average.

• Historical

- MAE (Lead Time)
- MAE (Years)
- C <u>RMSE (Lead Time)</u>
- C <u>RMSE (Year)</u>
- Skill Score (Lead Time)
- Skill Score (Years)
- O Percent Difference
- C POD Above Threshold
- POD Below Threshold
- C FAR Above Threshold
- <u>FAR Below Threshold</u>
- 🔍 <u>Rank Histoqram</u>
- Rank Histogram (Lead Time)
- Climate Variability
- C <u>Contingency Table</u>

Data Sources Time Scale Threshold

Data Source Information

Options

Statistic Data Sources

Coordinated Forecast

National Weather Service

Natural Resource

Conservation Service

Statistical Water Supply

California Department of Water

<u>Resources</u>

Ensemble Streamflow Prediction

Ensemble Streamflow Prediction - Empirical

Ensemble Streamflow Prediction - Normal

Ensemble Streamflow Prediction - Lognormal

Ensemble Streamflow Prediction - Wakeby

Ensemble Streamflow Prediction - Logwiebull

Ensemble Streamflow Prediction - Weibull

Prediction - Loglogistic

Time Scale

Threshold

Data Source Information

• Forecast Type

- Change Data Sources to be displayed
- Graph change on click

Options				
Statistic				
Data Sources				
Time Scale				
Period: Apr-Jul				
Years: 1996 ▲ 1997 1998 1999 2000 2001 2001 2002 ▼				
Month: Jan 💌				
Threshold				
Graph Options				
Data Source Information				

- Time Scale
 - Change Period
 - Modify Years
 - Graph change on click
 - Month
 - Only displayed when Contingency Table Statistic is chosen
 - Change the month the table displays



- Threshold
 - Default is Climatology / Historical Average
 - Enter value in KAF and press enter
 - Type 'mean' to return to default
 - Valid option for all but Rank Histogram statistic options.
 - Graph change on 'Enter'

Options				
	Statistic			
Da	ata Sources			
1	lime Scale			
	Threshold			
Data So	ource Information			
COORD	100%			
NWS	60%			
NRCS	60%			
SWS	60%			
CADWR	0%			
ESP	56.7%			

- Data Source Information
 - Calculated the percent of forecasts available
 - Keep in mind; can skew thoughts about statistics

Example

DIRC2 1991 – 2008 Forecast Available

Historical

Volume Forecast (kaf)

Historical - BLUE - DILLON RES



Forecast Period: Apr - Jul

Streamflow Histogram





- Graph Options
 - Any Combination of Options
 - Sort/Unsort Years
 - Display selected Years or All Observed Years
 - Display Summed Months
 - Individual Months
 - Each Month and Summed

Streamflow Histogram



Streamflow - BLUE - DILLON RES (DIRC2)

Forecast Period: Apr - Jul



Observed Year



Error and Skill By Lead Time

COORD NWS

NRCS SWS

ESP Threshold



Comparison of forecast error to • "average" error useful diagnostic tool



Graph Options

Display/Undisplay years within the months

Root Mean Squared Error Skill Score Relative to Threshold - 167 KAF BLUE - DILLON RES (DIRC2) Forecast Period: Apr - Jul



COORD
NWS
NRCS
S₩S
ESP

Error and Skill By Lead Time

Root Mean Squared Error - BLUE - DILLON RES (DIRC2)

Forecast Period: Apr - Jul



Error and Skill By Year



Options				
Statistic				
Data Sources				
Time Scale				
Threshold				
Graph Options				
Months: Undisplay 💌				
Data Source Information				

- Warning
 - SS can be miss leading when observed close to threshold (2003)

•

Graph Options

 Display All Months or Specific forecast Month





– C	OORD
📃 N	ws
🗖 N	RCS
🗖 🗖 S	ws ∣
Ε Ε	SP

Error and Skill By Year

Mean Absolute Error - BLUE - DILLON RES (DIRC2)

Forecast Period: Apr - Jul







Probability of Detection

- High years much for difficult to detect in the early season
- All forecasts during low years have been for low volumes



Probability of Detection - BLUE - DILLON RES (DIRC2) Volumes Below Threshold - 167 KAF Forecast Period: Apr - Jul



False Alarm Rate

ESP

 Similar story here as with POD



COORD

NWS

ESP





Forecast Distribution

- Some tendency to underforecast
- 26% of observed streamflow falls above the 10% exceedance forecast value
- Reasonable max not so reasonable
- Results very dependent on years selected



Forecast Uncertainty by month



Observed Lag-1 Analysis Climate Variability



years in database

Contingency Table

Contingency Table for Jan with Threshold - 167 KAF BLUE - DILLON RES (DIRC2)

COORD NWS NRCS SWS ESP		Observed Streamflow						
		Much Below < 125.25 KAF	Below 125.25 - 167 KAF	Above 167 - 208.75 KAF	Мисһ Аbove > 208.75 КАР			
F o r	Much Below < 125.25 KAF	2002	1991 2000 2000 2000 2000	1999 1999 1999 1999	1995 1995 1995 1995			
e c a s t S t	Below 125.25 - 167 KAF	1994 1998 2002 2004 2005 1994 1998 2002 2004 2005 1994 1998 2004 2005 1994 1998 2002 2004 2005 1994 1998 2002 2004 2005	1992 2001 2001 2001 1992 2001 1991 1992 2000 2001	1993 2003 1993 2003 1993 2003 1993 2003 1993 1999 2003	1995			
r e a m	Above 167 - 208.75 KAF			2007 2008 2008	1996 1997 1996 1 997 1996 1997 1996 1997			
I o W	Much Above > 208.75 KAF			2006	1996 1997			



• Graph Options

- Select the percentage to adjust the categories
- 25% Default
- Month Displayed
 January by default







DIRC2 Streamflow-Climate Variability Relationship



Results that might be presented ...

ESP is really, really good

- ESP is much better than any other forecast method particularly in high years
- With "real" ESP as opposed to ESP reforecasts, results should be even better



Forecast Quality

- January forecasts are essentially as good as climatology
- Coordination process appears to add (marginal) value except in April.
- Forecast tweaks much less than the error should not be entertained unless some overriding rational exists to support those tweaks.



High flows hard Low flows no so hard

- Forecast system is perfect (POD = 100% in all cases) for detecting below average flows.
- Forecasts struggle with detecting high flows even through May.



Reasonable max not so reasonable

 Observed streamflow greater than the reasonable max nearly 30% of the time.



Climate variability not much help

- Very low predictability based on prior year hydrology or climate index
- Weak tendency for low years to follow low years



Possible Application (Discussion)

- April DIRC2 forecast is 150kAc-ft. Average error for April is 35kAc-ft. How could you use this information to:
 - Improve your forecast?
 - Improve your forecast process?
 - Improve forecast application?

Great Basin Case Study credit: Brent Bernard

 Findings indicate we detect low flows better than high flows with current and often over forecast the 10% exceedance values.



Probability of Detection - LOGAN - LOGAN, NR, STATE DAM, ABV (LGNU1) Volumes Above Threshold - 126 KAF Forecast Period: Apr - Jul





Bear River – Near Utah, Wyoming State Line



Rank Histogram





 That many of the data sets from 1991-2008 have a binary split indicating that more extremes have occurred in observed flow during the last 20 years.



