HEFS workshop, 03/11/2015

Seminar B: hindcasting concepts and requirements

James Brown

james.brown@hydrosolved.com

Contents

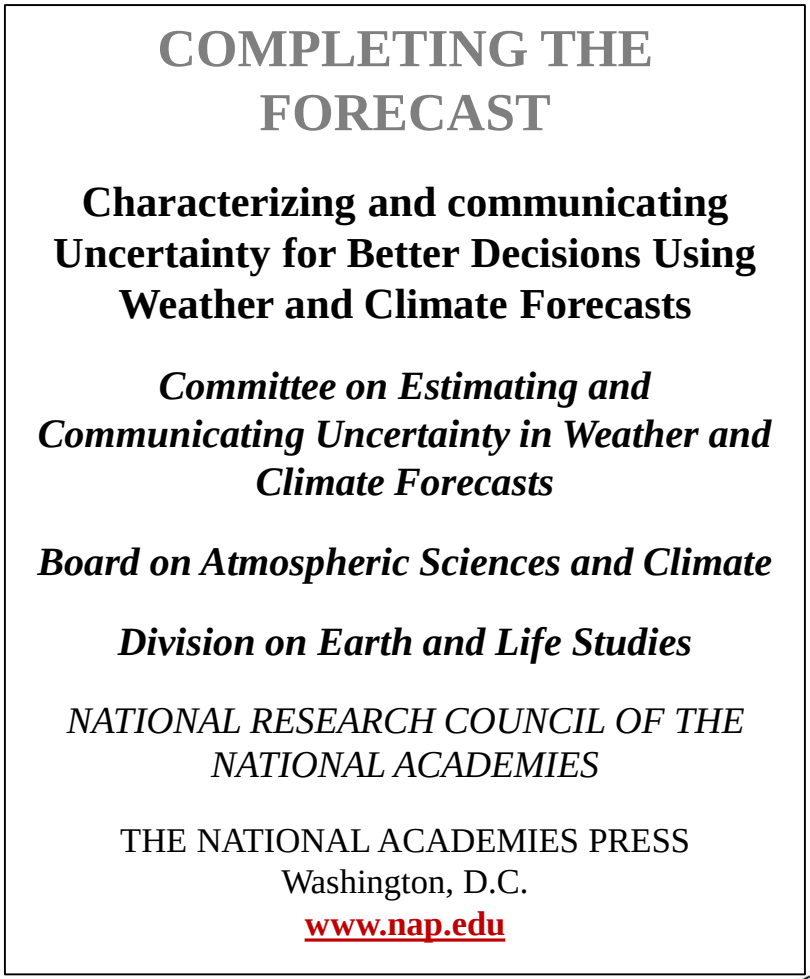
1. Why conduct hindcasting?
2. What are the data requirements?
3. How to design a validation study?
4. How to configure CHPS hindcasting?
5. Practical tips and considerations
6. Summary and final thoughts

1. Why conduct hindcasting?

Motivation: support verification

National Research Council, 2006

"Recommendation 6: NWS should expand verification of its uncertainty products and make this information easily available to all users in near real time. A variety of verification measures and approaches (measuring multiple aspects of forecast quality that are relevant for users) should be used to appropriately represent the complexity and dimensionality of the verification problem. Verification statistics should be computed for meaningful subsets of the forecasts (e.g. by season, region) and should be presented in formats that are understandable by forecast users. Archival verification information on probabilistic forecasts, including model-generated and objectively generated forecasts and verifying observations, should be accessible so users can produce their own evaluation of the forecasts."



Motivation: support verification

Hindcasting in support of verification

- Verification requires a large sample (including extremes)
- Verification requires a consistent sample
- Verification requires a relevant sample (i.e. current HEFS)

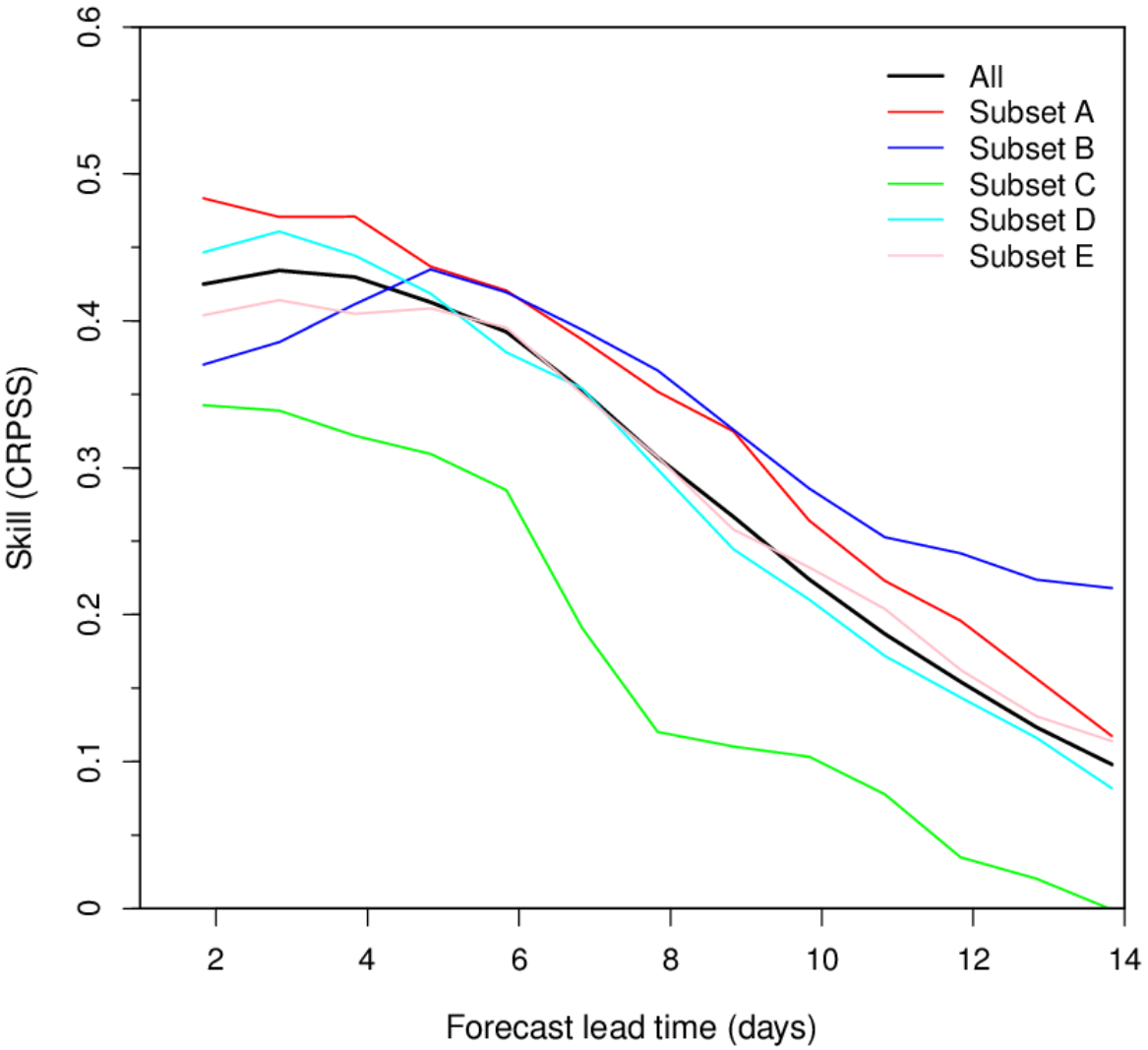
Verification in support of operations

- By leading to targeted improvements in the HEFS
- By improving guidance and building confidence in HEFS
- By providing historical analogs to forecast conditions
- By enriching forecast products (with quality information)
- By allowing end users to optimize their decision support

Motivation: example

Long/consistent archive

- Skill of HEFS hindcasts with MEFP-GEFS forcing vs. resampled climatology
- Fort Seward (FTSC1) in CNRFC
- Hindcasts for 15 years (1985-1999), also split into five, 3-year, sub-periods
- RFC QPF/QTF archives may be a few years only
- Using a short archive could give noisy and misleading results (leading to wrong conclusions)



2. What are the data requirements for hindcasting?

Data requirements

HEFS calibration data

- Hindcasting starts w/ operational setup/calibration
- Forcing: MAP/T/PE & raw forecasts (per source)
- Flow: QME/QINE and historical simulations

Other operational forecasting datasets

- Aim of hindcasting is to reproduce operations
- Archived diversions, extractions, releases?
- Other manual modifications archived?
- Probably not, but sensible to minimize differences

Data requirements

Sample size requirements

- Large and consistent historical sample needed
- Hindcasting improves length and consistency…
- … but basins and datasets still change over time
- Rule of thumb: 10+ years for verification

Verification needs vary with application

- Events of interest: threshold, seasons etc…
- …extremes (e.g. floods) need much longer record
- Other factors: forecast quality, basin memory

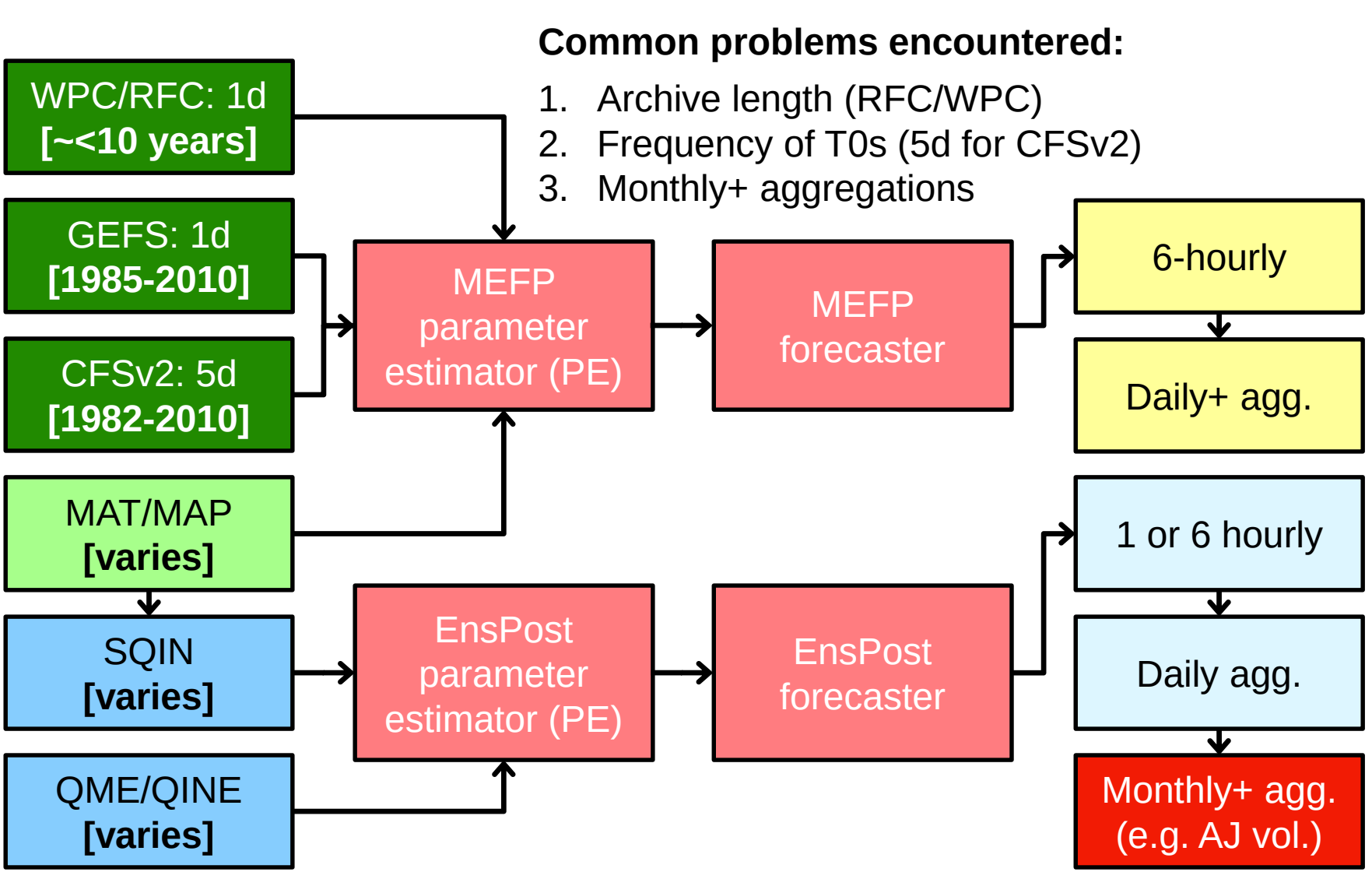
Impacts of threshold on sample size

Average sample sizes by threshold

Return period for daily mean flow	Expected sample size by reforecast type (N years, M days between T0s)								
	(30,1)	(25,1)	(20,1)	(15,1)	(10,1)	(25,3)	(25,5)	(25,7)	(10,5)
1 in 30 days	365	304	243	183	122	101	61	43	24
1 in 60 days	183	152	122	91	61	51	30	22	12
1 in 90 days	122	101	81	61	41	34	20	14	8
1 in 180 days	61	51	41	30	20	17	10	7	4
1 in 1 year	30	25	20	15	10	8	5	4	2
1 in 2 years	15	13	10	8	5	4	3	2	1
1 in 5 years	6	5	4	3	2	2	1	<1	<1
1 in 10 years	3	3	2	2	1	<1	1	<1	<1
1 in 20 years	2	1	1	<1	<1	<1	<1	<1	<1

- Red cells: events are nominally “unverifiable” (≤ 20 samples)
- Current GEFS reforecasts at edge for 1-2yr event (circled)
- Best case scenario: 20 samples may contain multi-day events

Controls on hindcast sample size



Mitigating sample size issues

Less than perfect archives are a reality

- For example, RFC forcing data may be <5 years
- Observed data may be missing or inadequate etc.

What steps can be taken?

- Will short archive add value (RFC vs. GEFS)?
- Use HEFS diagnostics to identify issues early
- For calibration, see MEFP sampling options
- For verification, focus on lumped scores, avoid extreme thresholds, assess sampling uncertainty

Mitigating consistency issues

Consistency issues are varied/complex

- Hindcasting removes many issues (v. archiving forecasts)
- But: do hindcasts cover consistent basin conditions?
- But: do hindcasts represent current operations?

What steps can be taken?

- Use latest CHPS/EnsPost; redo inconsistent hindcasts
- Archive mods and operational forecasts
- Check for inconsistencies and impacts (time consuming!)
  - Compare the archived operational HEFS forecasts & hindcasts
  - Examine changes in forecast quality over hindcast record

3. How to design a validation study?

Dependent/independent validation

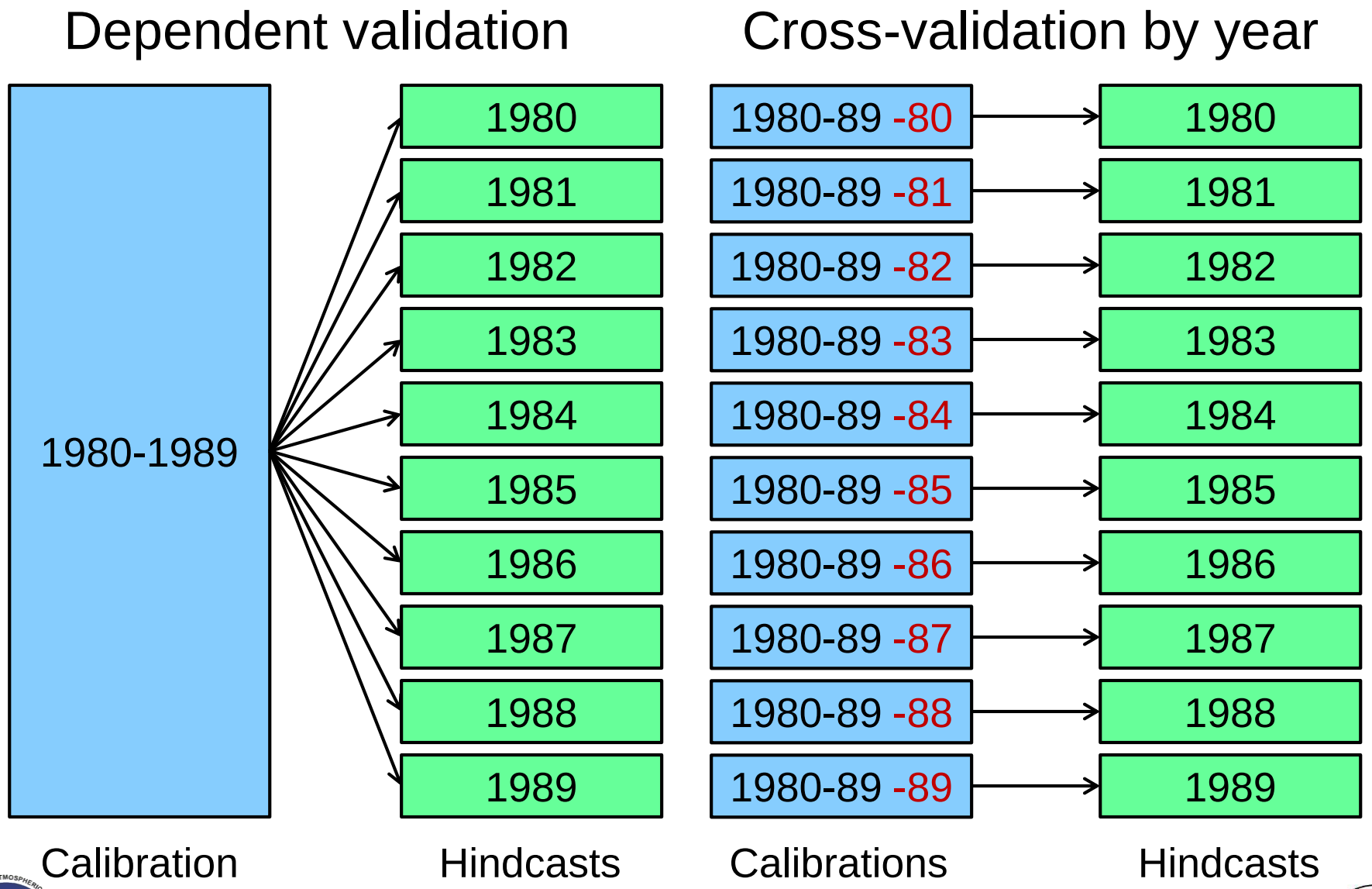
Dependent validation (practical)

- Calibration and validation periods are the same
- Advantage: simple, requires only one calibration/hindcast
- Disadvantage: exaggerates skill, particularly for extremes
- The approach used in the phased validation of the HEFS

Independent validation (preferable)

- Validation period does not overlap with calibration period
- Advantage: completely independent test of system
- Disadvantage: requires multiple calibrations/hindcast runs
- Several different flavors of independent validation…

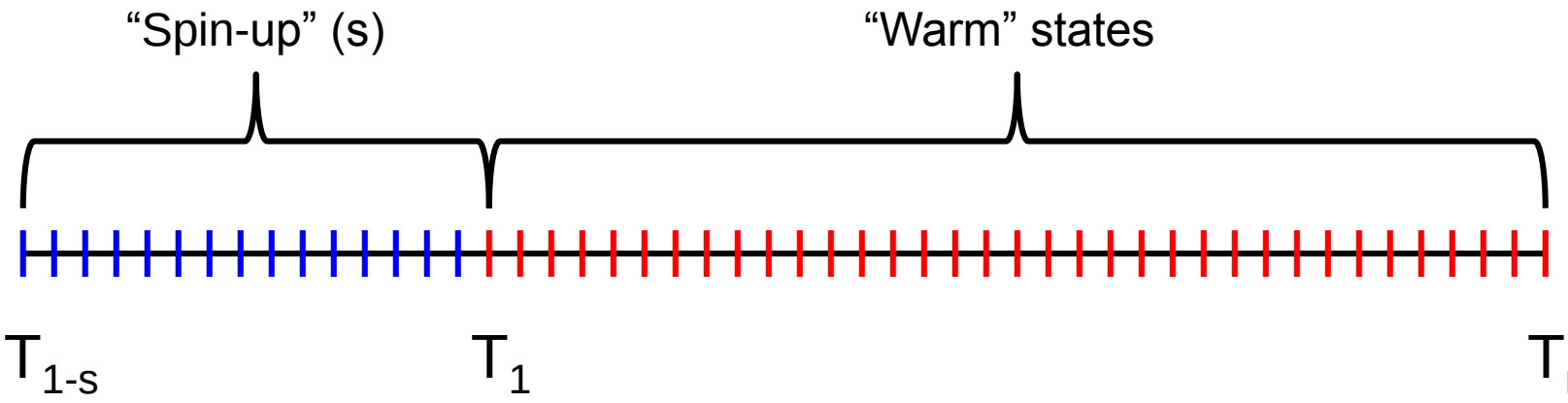
Examples of validation design



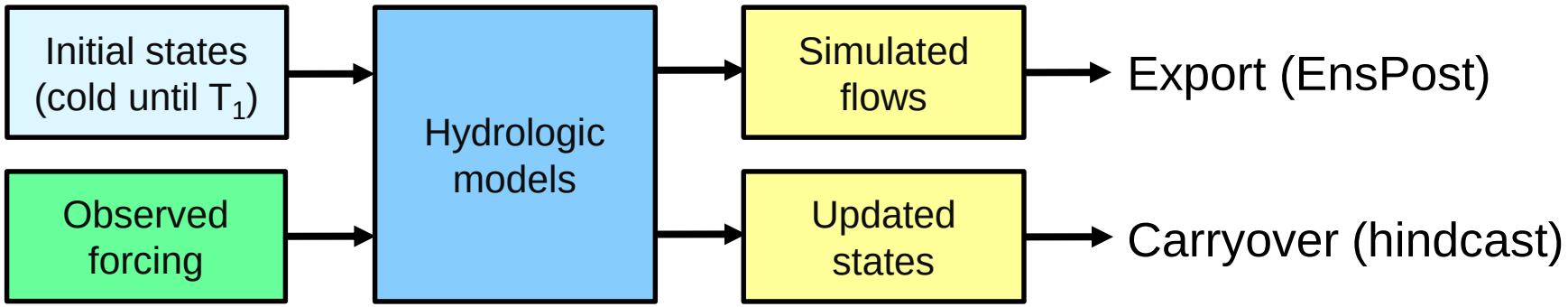
4. How to configure CHPS for hindcasting?

HEFS hindcasting mechanics

STEP 1: warm states and simulations for hindcast period

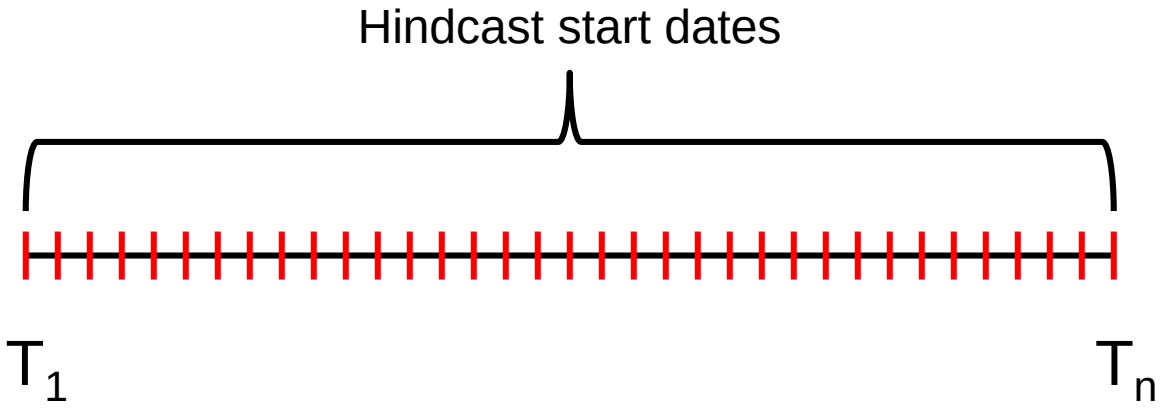


For each of T_{1-s} ,…,T_1,…,T_n:

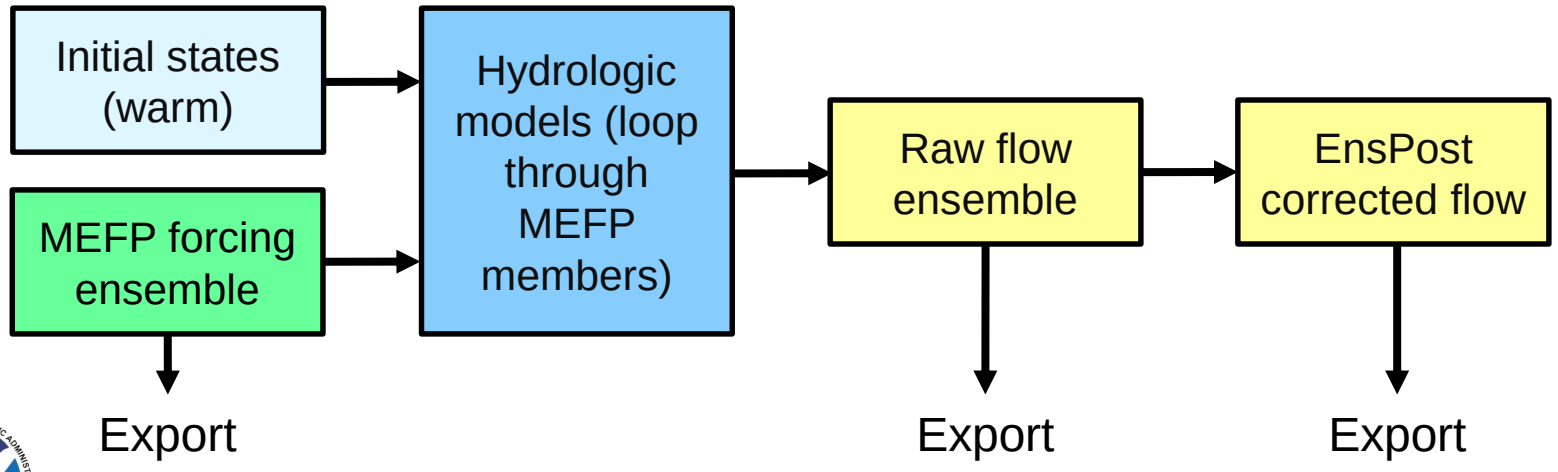


HEFS hindcasting mechanics

STEP 2: generate ensemble hindcasts



For each of $T_1,\ldots,T_n$:



Steps before generating hindcasts

- Entry point is a working operational HEFS in CHPS
  - MEFP calibrated and configured for operations
  - EnsPost calibrated and configured for operations
  - May need to adjust operational configs (tips later)
- Required data available for hindcast period
- Warm states generated (run “UpdateStates” first)
- Hindcasting configuration developed
  - Controls order of activities, such as running models & exporting
  - Configures the exports of the hindcasting datasets
  - Assumes dependent validation (single set of parameters)

Main steps (see Demo 1/Ex. 1)

Activity hierarchy (HEFS_Hindcast.xml)

1. Run MEFP P (EXISTS) →
2. Run MEFP T (EXISTS) →
3. Export MEFP **(NEW)** →
4. Run raw flow (EXISTS) →
5. Export raw flow **(NEW)** →
6. Run EnsPost (EXISTS) →
7. Export EnsPost **(NEW)** →

```
<activity>
    <runIndependent>false</runIndependent>
    <workflowId>MEFP_Temp_Forecast</workflowId>
</activity>
<activity>
    <runIndependent>false</runIndependent>
    <workflowId>MEFP_Precip_Forecast</workflowId>
</activity>
<activity>
    <runIndependent>false</runIndependent>
    <workflowId>MEFP_Export</workflowId>
</activity>
<activity>
    <runIndependent>false</runIndependent>
    <workflowId>HEFS_Forecast</workflowId>
</activity>
<activity>
    <runIndependent>false</runIndependent>
    <workflowId>HEFS_Export</workflowId>
</activity>
<activity>
    <runIndependent>false</runIndependent>
    <workflowId>EnsPost_Forecast</workflowId>
</activity>
<activity>
    <runIndependent>false</runIndependent>
    <workflowId>EnsPost_Export</workflowId>
</activity>
```

5. Practical considerations and lessons learned

Run times and disk space required

- Many factors will impact resources required
  - Number of forecast points
  - Forecast scenarios (e.g. climatology, flow w/ and w/o EnsPost,…)
  - Forecast time horizon and frequency of T0s
  - Hardware (run times)
  - Output formats (ASCII or compressed)
  - EVS outputs generated (pairs, compression, plots etc.)
  - Many others…
- Various hindcasting tests conducted at OHD…

Resources required

Example runs at OHD (hindcasts only)

Run property	Scenario 1	Scenario 2	Scenario 3
Forecast horizon (days)	15	15	365
RFC	MARFC	CNRFC	MARFC
Number of MEFP basins	14	28	14
Number of flow basins	14	15	14
Years (# years)	1988-1998 (10)	1985-1995 (10)	1988-1998 (10)
HEFS components	All (no G. Gen.)	All (no G. Gen.)	All (no G. Gen.)
Frequency of T0s (days)	1	1	5
Model timestep (hours)	6	1	6
Forcing sources	GEFS	GEFS	GEFS-CFSv2-CLIM
Runtime per T0 (mins)	0.75	1.4	3.2
Runtime per year (mins)	278	517	240
Total run time (mins)	2780	5170	2400
MEFP as % of run time	32	60	-
localDataStore (GB)	16	38	41
PI-XML total export (GB)	43	71	14.5

General tips (not RFC specific)

- Data QC
  - Use test runs (e.g. 2 yr) to screen for obvious issues
  - Check exports created for each T0 (or use Hindcasting Robot)
  - Search (e.g. grep) for missing data in export files
  - Basic QC before verification; verification before application!
- Manage disk-space requirements
  - Plan for disk-space requirements before run
  - Set time series as “temporary” (targeted but cumbersome)
  - Reduce exports to CHPS log (log.txt); this can grow to many GB
  - Can export compressed (.fi/.bin) files, but not human readable

General tips (not RFC specific) cont.

- Manage runtimes
  - Configure CHPS for parallel processing (FogBugz #1150)
  - Can split runs manually or use Hindcasting Robot
  - Run on a local disk (about 30% faster)
  - Restrict workflow to locations of interest (avoids searching)
- Manage/avoid runtime failures
  - Create virtual “sandpit” for hindcasting with vncserver/vncviewer
  - If not using vncviewer, turn off screen saver to avoid freezing
  - Break CFSv2 runs around Feb 29th (skip) to retain 5-day cycle
  - Avoid timeout by increasing runtime limit (e.g. 600 secs per T0)
  - In DB viewer, use F12 + M to terminate run

General tips (not RFC specific) cont.

- Export considerations
  - Set HEFS modules runIndependent="false". This avoids silent exporting of legacy (incorrect) data from earlier runs
  - If possible, export files per basin/variable (easier to verify)
  - pi-xml is easier to work with, .fi/.bin is smaller/faster for I/O
- Finally, use the HEFS Hindcasting Guide
  - All these tips and more can be found in the Hindcasting Guide
  - Configuration and run checklists provided
  - Brief introduction to Hindcasting Robot
  - Separate manual also available for Hindcasting Robot

Tips that vary with RFC

- Warm states search window (end at 0): e.g. MARFC
- Data import/merge considerations
  - Import 6-hourly MAP/MAT etc. to closest time in (0Z,6Z,12Z,18Z)
  - MERGETS with MAPX priority but no MAPX = zeroes
  - Ensemble MAPE (set to “read all forecast”): e.g. MARFC
- Hindcast configurations out of sync with operations
  - Outdated config may conflict w/ latest binaries (e.g. LagK)
  - Syncing avoids this, but raises other issues
  - For example, EnsPost may require re-calibration

6. Summary and final thoughts

Final thoughts

Create checklist (Hindcasting Guide)

- Think about length and consistency of historical record
- Think about relevance for operational practice
- Identify scenarios needed (including baseline runs)
- Choose an experimental design (dependent validation)
- Adapt operational configs. for hindcasting if needed
- Generate warm states and simulations (for EnsPost)
- Optimize run settings (parallel process, split runs etc.)
- QC input data/parameters and conduct test runs (ideally)
- QC final runs (ideally) and conduct verification

Questions?

Extra slides

Runtime modifications

Hindcasts vs. operational forecasts

- Many adjustments made in real-time are not archived
- Thus, hindcasts will differ from operational forecasts
- If possible, compare hindcasts & (archived) forecasts
- How to minimize runtime mods and archive others?

Calibration vs. operational forecasts

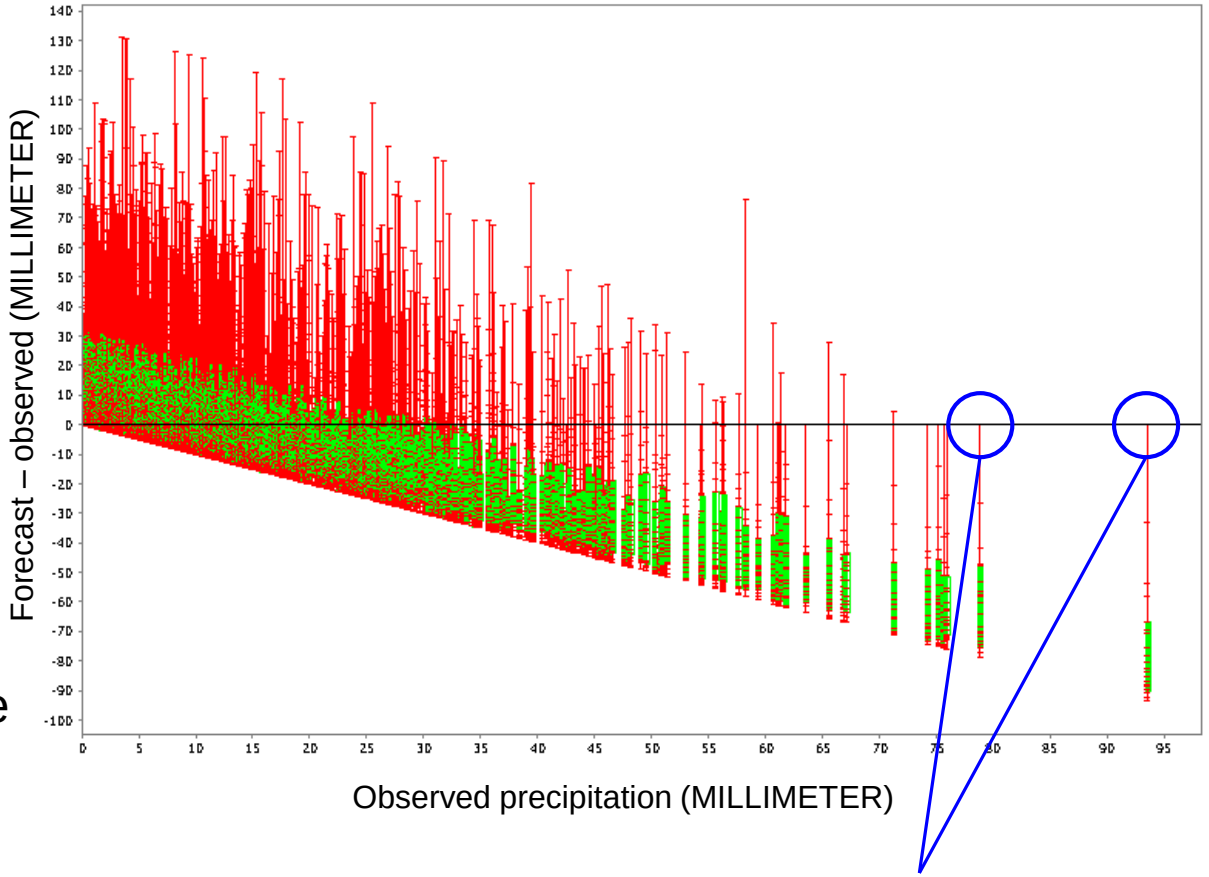
- EnsPost: need consistent historical & operational sim.
- If operational simulations differ, EnsPost not optimal
- Again, how to minimize differences or archive mods?

Warning: MEFP raw climate option

Dependent validation

- Example of problems with dependent validation
- MEFP "raw climatology" samples historical observations without fitting/smoothing
- In hindcast mode, one ensemble member is always equal to the verifying observation!
- Not realistic, so do not use MEFP raw climatology in hindcast mode
- Otherwise, dependent validation still "best" option

MEFP raw climatology, precipitation at WALN6



In dependent validation, one member in raw climatology always equals verifying observation!