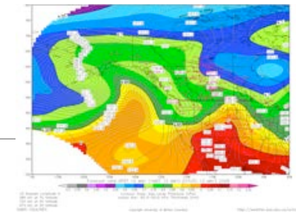
Terrain Issues, Post-processing, & Operational Forecasting

Roland Stull
Geophysical Disaster Computational Fluid Dynamics Center
Earth, Ocean & Atmos. Science Dept.
University of British Columbia
2020-2207 Main mall
Vancouver, BC V6T 1Z4 Canada

rstull @ eoas.ubc.ca

Special Thanks to my NWP Team: Henryk Modzelewski, George Hicks II, Thomas Nipen, Dominique Bourdin, Atoossa Bakhshaii, May Wong, Bruce Thomson, Greg West, Kyle Sha & many former team members.

Topics



- •Terrain Issues (from previous lecture)
- •Post-processing
- •Operational Forecasting
- •NWP Course Summary

Terrain Issues

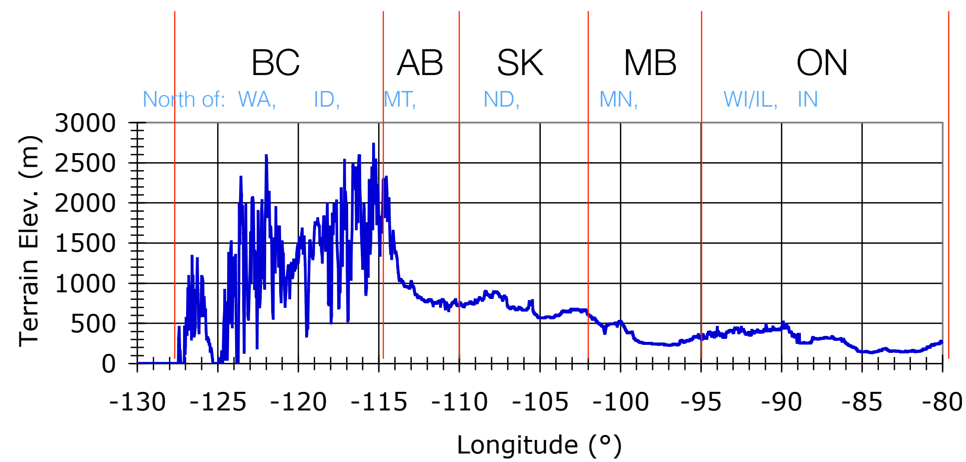
Modeled terrain is smoothed relative to actual terrain.

Mountain tops are cut off, and valleys are filled in. More so for coarser grids.

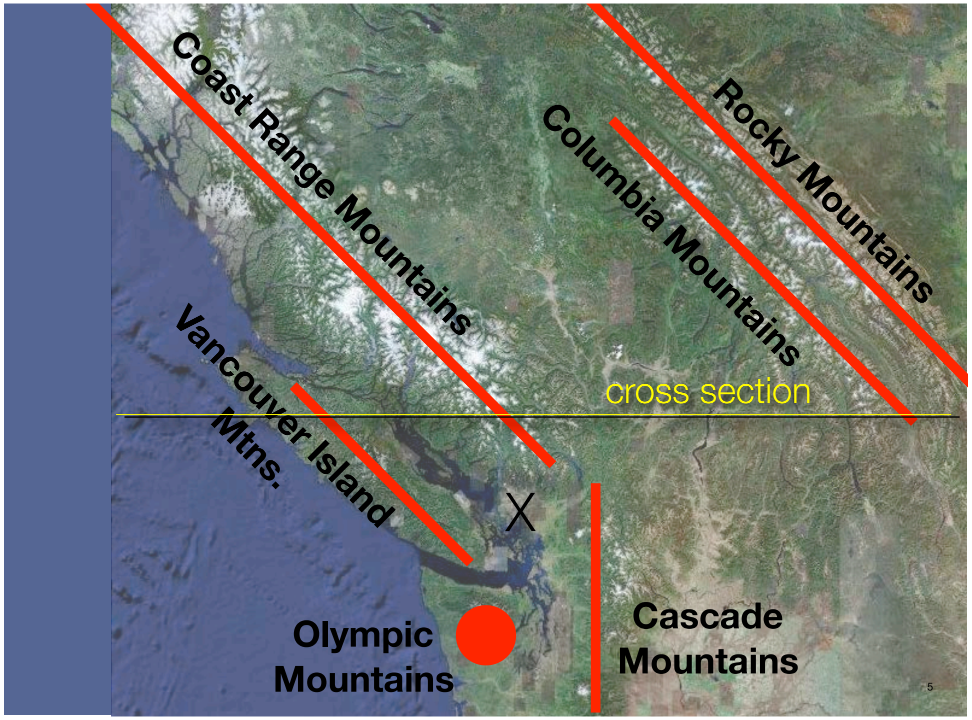
Thus, the elevation of a verifying obs is often different from the elevation in the NWP model.

Thus, the model is **not representative** of the observation.

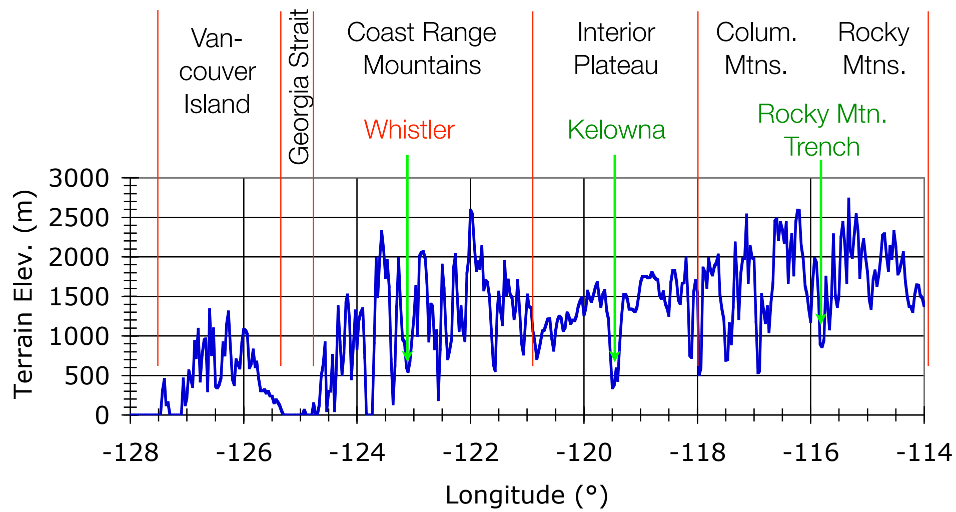
Canadian Terrain Elevation



- West-East terrain cross section through Whistler (50.12°N)

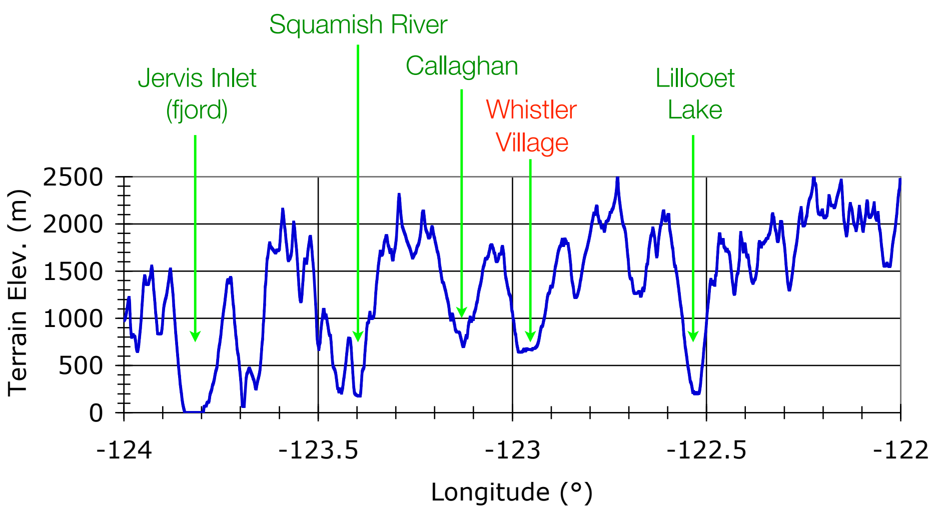


British Columbia Terrain Elevation



- West-East terrain cross section through Whistler (50.12°N)

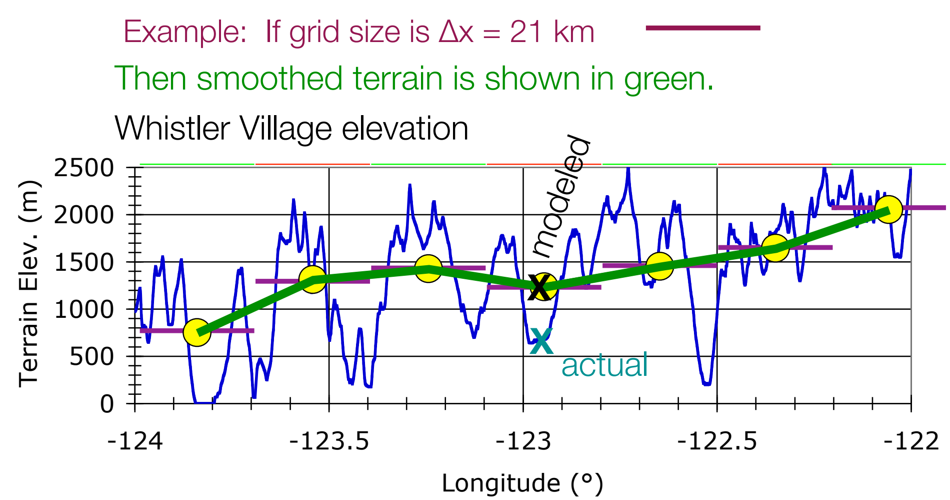
Zooming in Near Whistler



- West-East terrain cross section through Whistler (50.12°N)

Terrain Elevation vs. Grid Size.

Terrain must be smoothed to match grid resolution.



- West-East terrain cross section through Whistler (50.12°N), where 0.1°lon ≈ 7 km.

How Fine is Fine Enough?
Many valleys are narrower than 1 km

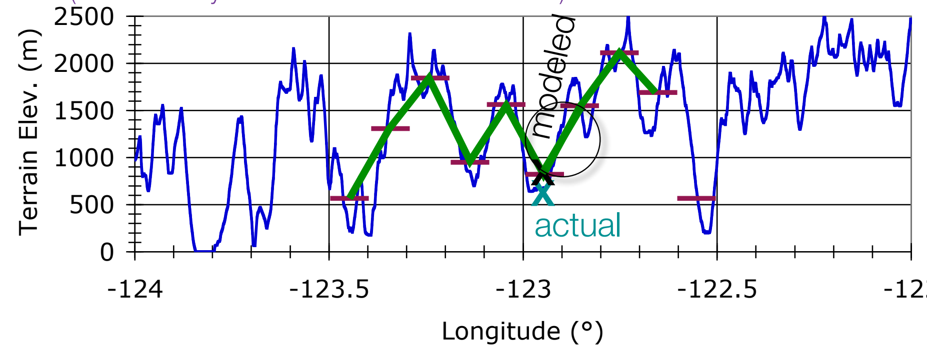
An Obvious Trick: Use finer horizontal grid size

Example: If grid size is Δx = 7 km

Then the modeled terrain is closer to the actual terrain. **Good.**

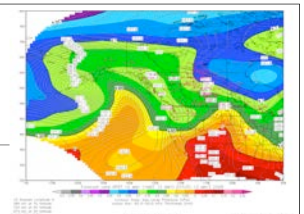
And the modeled slopes become steeper (closer to real). **Difficult**. Violate CFL.

(& still have systematic errors due to location.)



- West-East terrain cross section through Whistler (50.12°N), where 0.1°lon ≈ 7 km.

Topics



- Terrain Issues (from previous lecture)
- Post-processing
- Operational Forecasting
- NWP Course Summary

Post-Processing of Deterministic Forecasts
General procedure:

- Compare past forecasts with past observations (bias = fcst – obs)
- Compute the average past bias
- Assume the future bias is equal to (or related to) the average past bias.
- Add this bias correction to the raw NWP forecast.
- Notes: different biases at
  - Different times of day
  - Different fcst horizon
  - Different locations
  - Different seasons
  - Different climate cycles
- Compensates for imperfections in the NWP dynamics, numerics, physics, terrain smoothing, location errors, etc.

Temperature Mean Error (Bias Estimate)

Mean Error Legend	<-3°	<-2°	<-1°	>1°	>2°	>3°

Station		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8
YVR	MINT24	0.0	0.1	0.1	0.1	0.0	0.1	0.1	0.2
	MAXT24	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2
COQ	MINT24	-0.1	0.0	-0.0	0.0	-0.1	0.0	-0.0	0.1
	MAXT24	-0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0
CQM	MINT24	-0.1	-0.0	-0.0	-0.0	-0.1	0.0	-0.0	0.2
	MAXT24	0.0	0.1	0.1	0.1	0.1	0.2	0.1	0.1
YXX	MINT24	0.1	0.2	0.1	0.2	0.1	0.2	0.1	0.2
	MAXT24	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.1

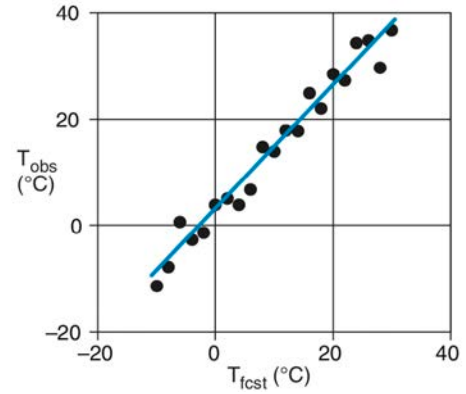
Post-Processing of Deterministic Forecasts
Terminology:

- **Predictand**: output = what you are trying to predict. E.g. temperature bias
- **Predictor**: input = what you know could be a factor

- **Training** data set: used to create the statistics or to find the regression
- **Testing** data set: independent data used to validate or select the best of several solutions (to avoid **over-fitting**: fitting the noise as well as the data)
- **Verification** data set: another independent data set to show the goodness/badness of the resulting bias-correction algorithm.

Post-Processing of Deterministic Forecasts
Methods

- **Perfect Prog**
- **MOS (Model Output Statistics)**
- **UMOS (Updatable MOS)**
- **Short-update-period MOS**
- **Kalman filters**
- **Gene-expression programming**
- **Artificial Neural networks**



Post-Processing of Deterministic Forecasts
Perfect Prog

- Used to for variables that the model doesn't predict directly
- Not used as much any more, because models predict more things
- "perfect" because regresses the pedictand against the observed (not fcst) predictor.

- Advantages: not dependent on any one model
- Disadvantages: doesn't correct for model errors; requires large data set

Post-Processing of Deterministic Forecasts
MOS

- MOS = Model output statistics
- Use 2 or more years of data.
  - Longer data set better captures rare events (extreme rains, winds, etc.)
  - Longer data set better captures climate cycles

- Advantages: corrects for NWP model imperfections
- Disadvantages: needs to be recomputed each time the NWP model is hanged; requires large data set; requires re-forecast of past events to create the 2-years of forecasts.

Post-Processing of Deterministic Forecasts
UMOS

- UMOS = Updatable model output statistics
- Same as MOS, except when a NWP model changes, then
  - Bias correction =
    (1 – w) · (Old MOS correction) + w · (New MOS correct.)
  - where weight w ≈ (Days Running New NWP) / (365 days)
- Advantages: like MOS, plus adapts faster to NWP model changes.
- Disadvantages: requires large data set for the old NWP MOS.

Post-Processing of Deterministic Forecasts
supMOS

- supMOS = short-update-period model output statistics
- Bias correction computed via a running average of recent past days
  - Today's correction = average bias from past few weeks
  - Options: Weighted averages (e.g., give more weight to recent biases).
    Different averaging times for different forecast horizons.
    - use past 5 days for forecast horizon of 1 day,
    - use past 3 weeks for forecast horizon of 8 days.
- Advantages: adapts faster to NWP model changes; adapts faster to synoptic regime and seasonal changes; needs a small data set
- Disadvantages: Misses extreme (rare) events.

Post-Processing of Deterministic Forecasts
KF

- KF = Kalman filter
- Recursive: uses today's bias to update overall bias correction from yesterday
- Predictor - corrector
- Advantages: adapts extremely fast to weather & model changes; needs to save only a few data; handles noisy data; provides "optimum" correction
- Disadvantages: Misses extreme (rare) events; misses synoptic regime changes. If incorrect Kalman gain is used, then gives too much weight to the most recent bias.

Post-Processing of Deterministic Forecasts
KF

INFO • Kalman Filter (KF)

Rudolf Kalman suggested a method that we can modify to estimate the bias x in tomorrow's forecast. It uses the observed bias y in today's forecast, and also uses yesterday's estimate for today's bias x_{old}:

$$x = x_{old} + \beta \cdot (y - x_{old})$$

The **Kalman gain** β depends on ratio $r = \sigma^2_{PL}/\sigma^2_{NWP}$, where σ^2_{PL} is the "predictability-limit" error variance associated with the chaotic nature of a "perfect" weather-forecast model, and σ^2_{NWP} is the error variance of the operational NWP model. If those error variances are steady, then $\beta = 0.5 \cdot [(r^2+4r)^{1/2} - r]$. The e-folding response time (days) is $\tau = -1/[\ln(1-\beta)]$.

Midlatitude weather is more variable and less predictable in winter. As a result, useful values are:
- Winter: $r \approx 0.06$, $\beta = 0.217$, $\tau = 4$ days.
- Summer: $r \approx 0.02$, $\beta = 0.132$, $\tau = 7$ days.

This Fig. shows a noisy input y (thin tan line) & KF responses x (thicker lines) for different values of the ratio r. The KF adapts to changes, and is recursive.

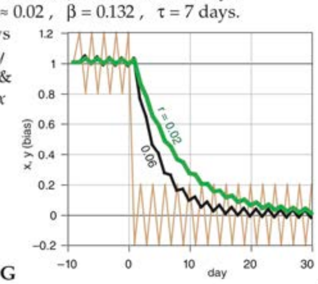
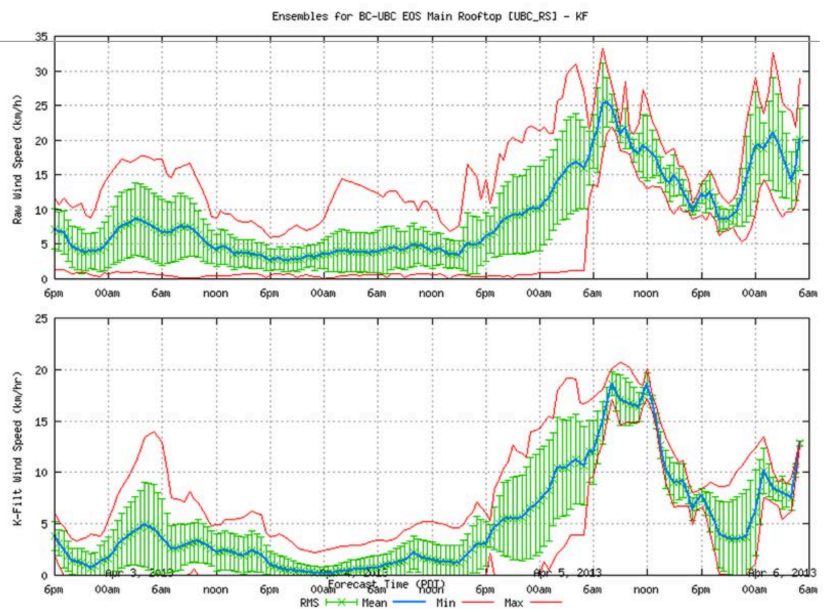
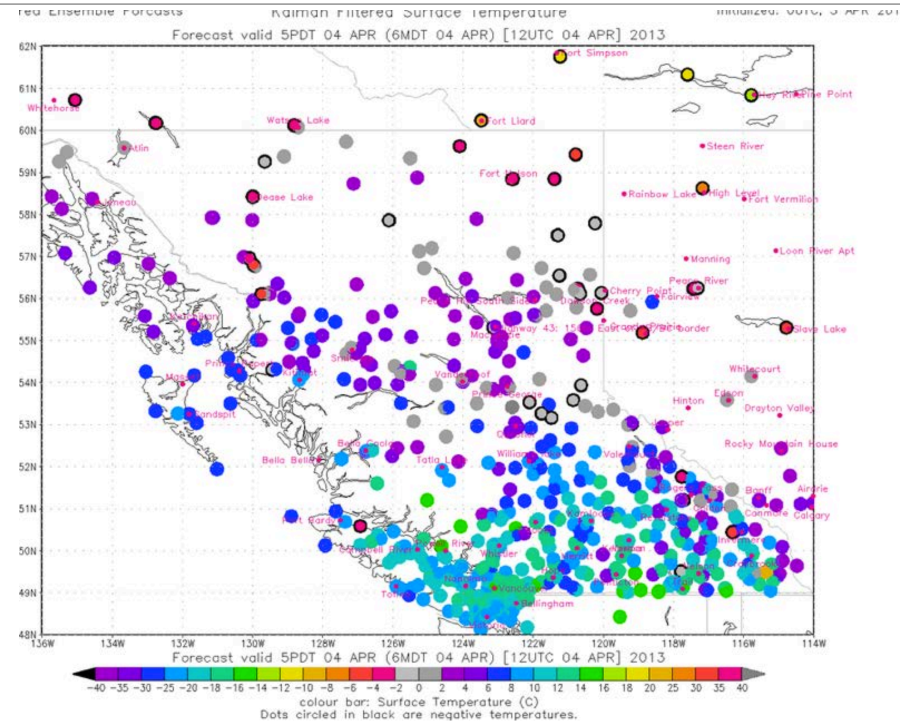


Fig. 20.G

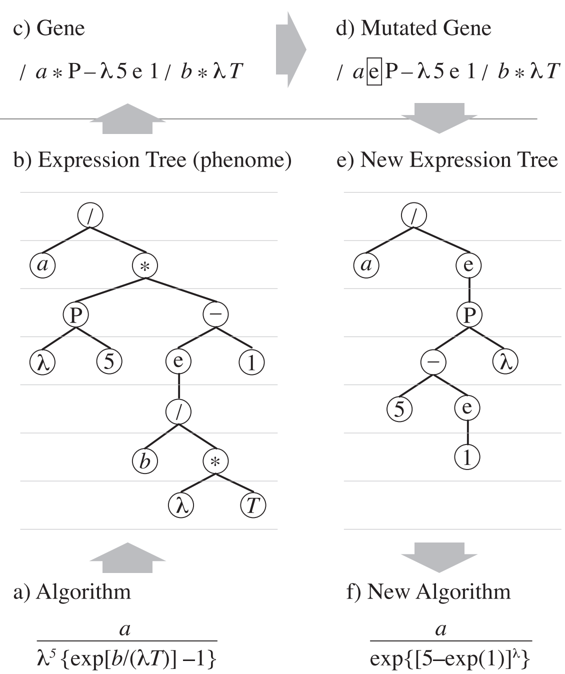
Post-Processing of Deterministic Forecasts
KF



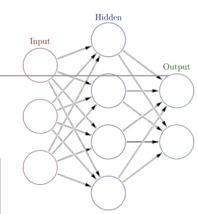


Post-Processing
GEP

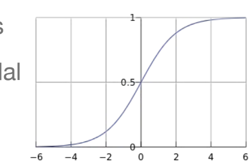
- GEP = Gene Expression Programming (evolution of algorithms)
- GEP regresses constants (weights) AND functions.
- We've used GEP to create ensemble averages, to estimate electric load, to approximate moist adiabats on thermo diagram, & to find Tw from T and RH.
- Advantages: can find complicated nonlinear relationships
- Disadvantages: Requires complex algorithms to optimize.



Post-Processing of Deterministic Forecasts
ANN



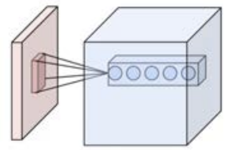
- ANN = Artificial Neural Network
- mimics connections between neurons
- operates on input data with a sigmoidal function such as hyperbolic tangent, error function, or logistic function.



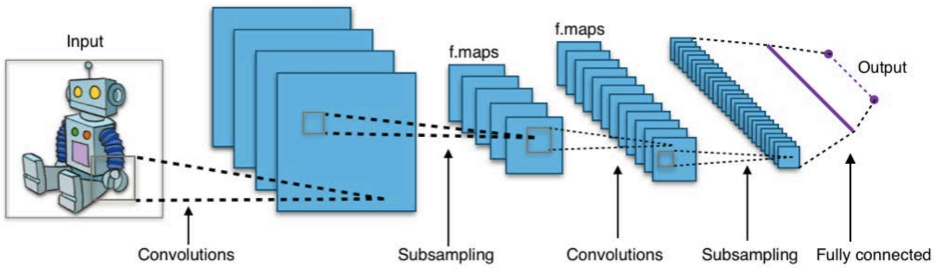
- uses a linear transfer function to combine hidden node results into output
- ANN regresses many 10s of weights for a few fixed functions. Example: we used ANN to predict electric loads for BC Hydro. Had 6 input nodes (corresponding to different weather and calendar variables), 10 hidden nodes, and 1 output node (electric load). This ANN had 81 free parameters.
- Advantages: can find complicated nonlinear relationships
- Disadvantages: Requires complex algorithms (e.g., feed forward, back propagation) to find the best weights; often overfits (fits the noise)

Post-Processing of Deterministic Forecasts
CNN

- CNN = Convolutional Neural Network (a deep learning ANN)
- thousands of free parameters (weights)
- used for classifying images (including weather maps)



Neurons of a convolutional layer (blue), connected to their receptive field (red)



https://en.wikipedia.org/wiki/Convolutional_neural_network

Post-Processing of Deterministic Forecasts
CNN

- Can be used for:
- Downscaling coarse-res NWP output to finer-res. fcsts.
- Quality Control
- Bias Correction
- This example from Kyle Sha, for downscaling.

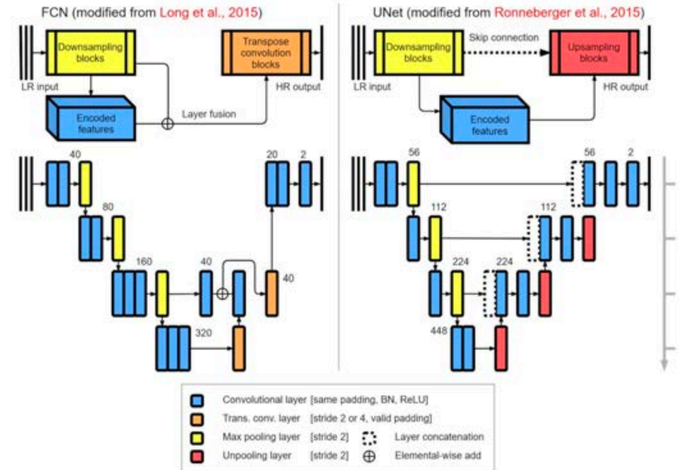


Figure 6. The architectures of two candidate CNN designs for temperature and precipitation downscaling. Numbers above the building blocks are the number of convolution channels. Vertical axis on the right shows the downsampling rate from 0x (no downsampling) to 8x (8 times downsampling).

See Kyle Sha's presentation at NCAR (Dec 2019): https://www.youtube.com/watch?v=mz7Qi-Hy7ho

Post-Processing of Deterministic Forecasts
CNN

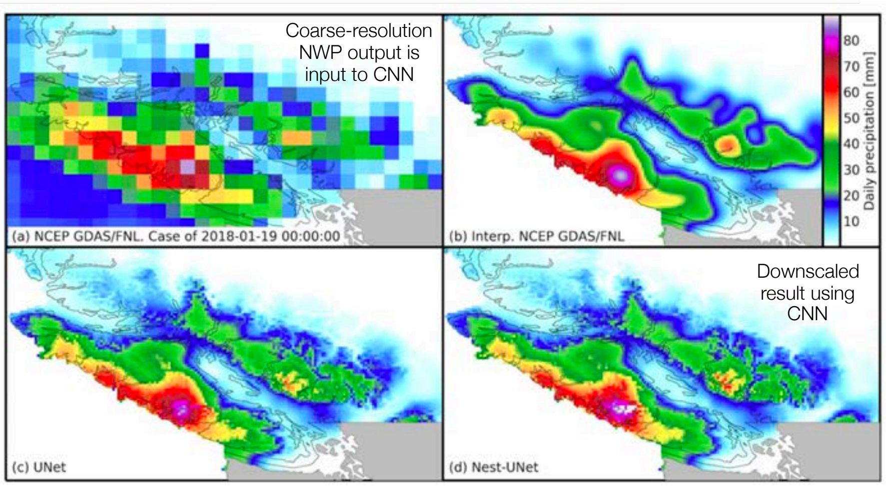
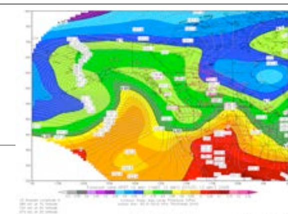


Figure 5. Daily precipitation downscaling example on 2018-01-19 and over British Columbia. (a) is the NCEP GDAS/FNL input, (b) is the interpolated FNL with ocean grid points masked out. (c, d) are the downscaling outputs of UNet and Nest-UNet.

Topics



- Terrain Issues (from previous lecture)
- Post-processing
- Operational Forecasting
- NWP Course Summary

Operational NWP
Overview

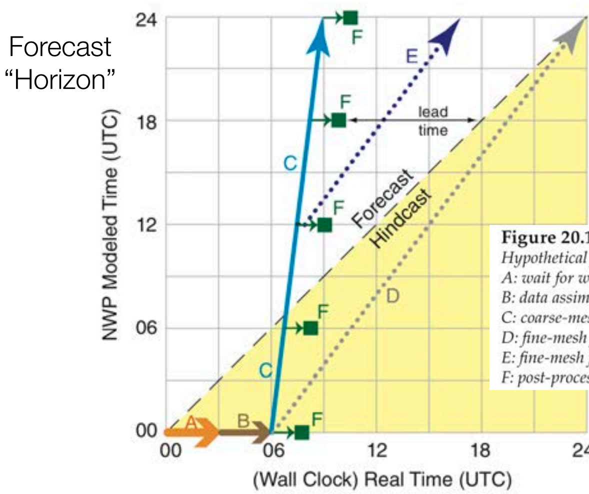
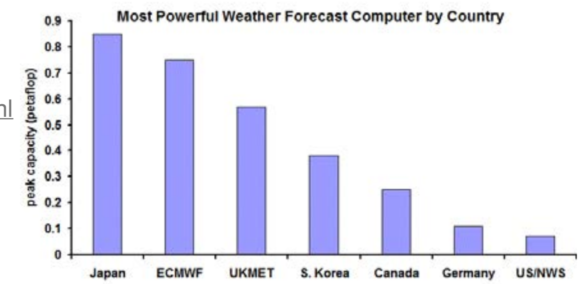


Figure 20.12
Hypothetical forecast schedule, for a 00 UTC initialization.
A: wait for weather observations to arrive.
B: data assimilation to produce the analysis (ICs).
C: coarse-mesh forecast.
D: fine-mesh forecast, initialized from 00 UTC.
E: fine-mesh forecast initialized from coarse forecast at 12 h.
F: post-processing and creation of products (e.g., weather maps).

Operational NWP
Model run status at forecast centers

- See "status" page on UBC web site http://weather.eos.ubc.ca/wxfcst/ (and give handouts)
- See "current status" under Environmental Models for NCEP web site http://www.nco.ncep.noaa.gov/
- Computers and their utilization. Sequencing. cron jobs. Time durations for input, fcst, graphics production.
- AT UBC we have 416 cores in our NWP computer clusters.
- See computer size issues at http://cliffmass.blogspot.ca/2013/02/the-us-weather-prediction-computer-gap.html



Operational NWP
Model run time vs. computer power

INFO • Moore's Law & Forecast Skill

Gordon E. Moore co-founded the **integrated-circuit** (computer-chip) manufacturer Intel. In 1965 he reported that the maximum number of transistors that were able to be inexpensively manufactured on integrated circuits had doubled every year. He predicted that this trend would continue for another decade.

Since 1970, the rate slowed to about a doubling every two years. This trend, known as **Moore's Law**, has continued for over 4 decades.

Increasing computer power has enabled improved NWP models that use finer grids covering larger areas with more-complicated physics and numerics. Thus NWP forecast skill has improved concomitant with computer power.

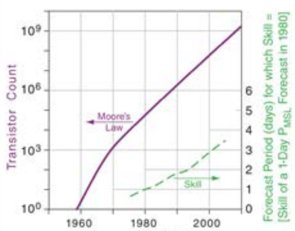


Fig. 20.D. *Moore's Law and forecast skill vs. time.*

INFO • Amdahl's Law

Computer architect Gene Amdahl described the overall speedup factor S_{ALL} of a computer program as a function of the speedup S_i of individual subroutines, where P_i is the portion of the total computation done by subroutine i:

$$S_{ALL} = \left[\sum (P_i / S_i)\right]^{-1}$$

and where $\Sigma P_i = 1$.

Special programs called **profilers** can find how much time it takes to run each component of an NWP model, such as for the implementation of the Weather Research & Forecast (WRF) model shown below.

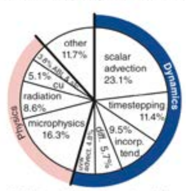


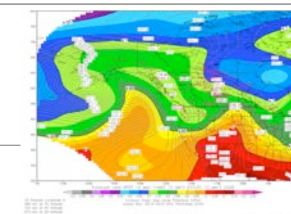
Fig. 20.F. *Portion of total run time of the WRF model for some of the major components.*
- *"incorp. tend." = incorporation of tendencies.*
- *"diff." = diffusion.*
- *"uvw advect." = advection of U, V, and W wind components.*
- *"microphysics" = hydrometeor parameterizations.*
- *"cu" = parameterizations for convective clouds.*
- *"ABL & sfc." = boundary layer and surface parameterizations.*

For example, if **graphics-processing units (GPUs)** speed up the microphysics 20 times and speed up scalar advection by 1.8 times (i.e., an 80% speedup), and the remaining 60.6% of WRF has no speedup, then overall:

$$S_{ALL} = [0.163/20 + 0.231/1.8 + 0.606/1]^{-1} = \mathbf{1.35}$$

Namely, even though the microphysics portion of the model is sped up 2000%, the overall speedup of WRF is 35% in this hypothetical example.

Topics



- Terrain Issues (from previous lecture)
- Postprocessing
- Operational Forecasting
- NWP Course Summary

ATSC 507 NWP - Course Overview
Topics 2020

1. Scientific basis for NWP (governing eqs)
2. Approximations to the governing eqs. Flux (conservative) forms.
3. Vertical coordinate transformations (terrain following, sigma)
4. The WRF model (Tim Chui). Vapor display software (Nadya Moisseeva)
5. Horizontal coordinate transformations (map projections, map factors)
6. Finite-difference methods: spatial and time. Acoustic split time diff. (Chui)
7. Finite-difference errors: truncation, amplitude, phase, group, nonlinear.
8. Finite-volume methods. MPAS & FV3. (Chui)
9. Spectral methods
10. Ensemble methods, ensemble avg & spread. Nonlinear dynamics & chaos.
11. Verification methods
12. Probabilistic forecasting (Thomas Nipen)
13. Post-processing methods & Operational forecasting

ATSC 507 NWP - Course Overview
Methods & Activities 2020

1. Textbook readings
2. Journal paper readings
3. Derivations, including by students at the blackboard
4. Understanding WRF (not a black box)
5. Running WRF & visualizing the output
6. Student presentations in class of WRF physics
7. What-if demos using spreadsheets and animations
8. Samples of code
9. Homeworks
10. Projects
11. Crunching numbers
12. In-class lab work (graphical interpretation of time diff. schemes)
13. Guest lectures (Tim Chui, Nadya Moisseeva, Thomas Nipen, Pedro Odon)
14. Flexibility in the computational tools you could use for your HW

ATSC 507 NWP - Course Overview
Conclusion

I learned a lot. Hope you did too.

-the end-