Today's Lab, Making Appendix 2: Data Rationalization

Transforming raw WxProj data into the final dataset….

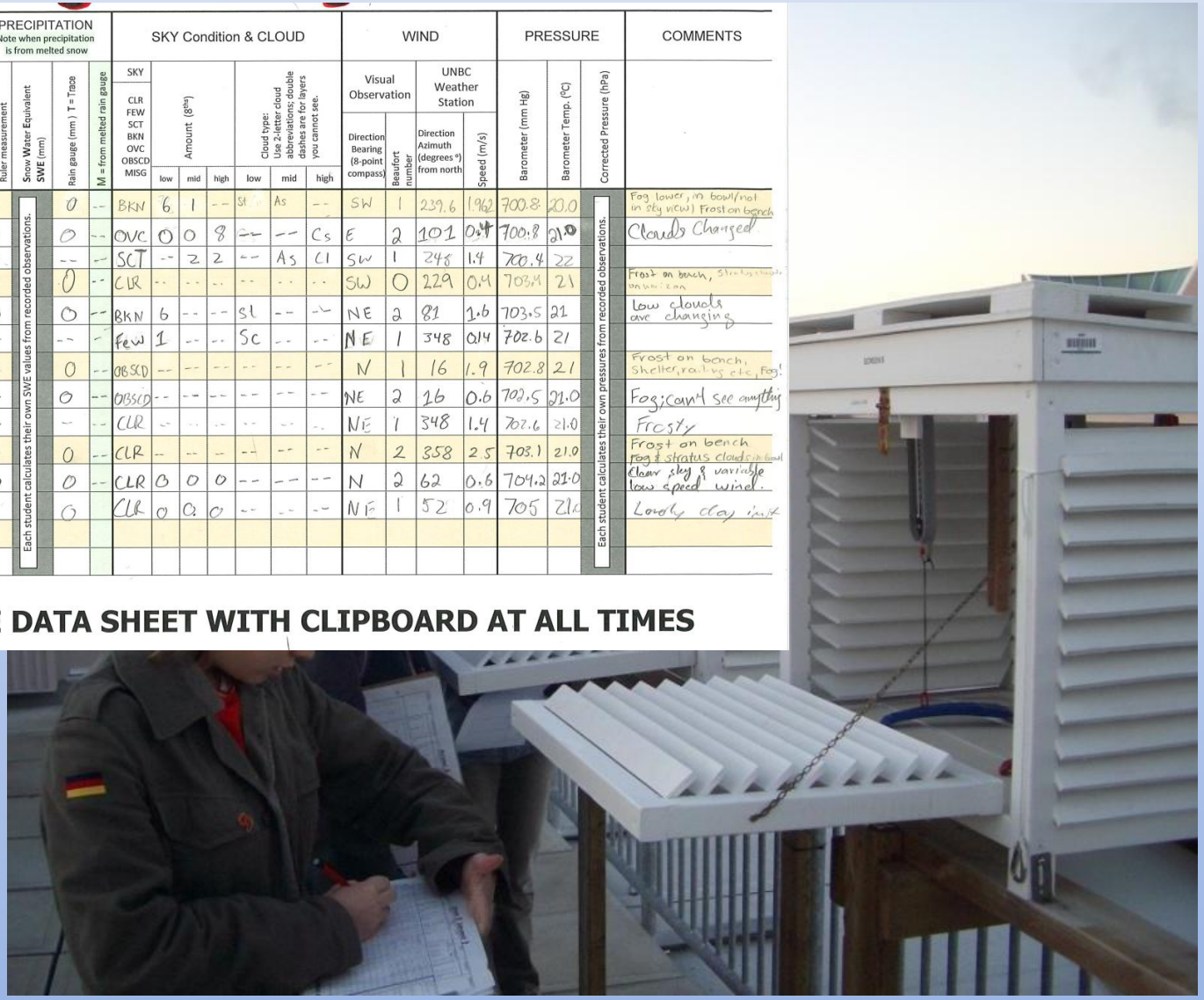
DATE	TIME	TEMPERATURE			HUMIDITY (Note wick state & instrument type in e & RH column)				PRECIPITATION (Note when precipitation is from melted snow)				SKY Condition & CLOUD								WIND				PRESSURE			COMMENTS
Entered from earliest to latest date & time		Max Temp. (°C)	Min Temp. (°C)	Present Air Temp. (°C)	Wet bulb Temp. (°C)	F = wick is frozen	Dry bulb Temp. (°C)	e (hPa) %RH A=Assmann B=Bacharach W=Weksler	Snow Depth (mm) Ruler measurement	Snow Water Equivalent SWE (mm)	Rain gauge (mm) T = Trace	M = from melted rain gauge	SKY: CLR FEW SCT BKN OVC OBSCD MISG	Amount (8ths) low	Amount (8ths) mid	Amount (8ths) high	Cloud type: Use 2-letter cloud abbreviations; double dashes are for layers you cannot see. low	Cloud type mid	Cloud type high	Visual Observation: Direction Bearing (8-point compass)	Visual Observation: Beaufort number	UNBC Weather Station: Direction Azimuth (degrees °) from north	UNBC Weather Station: Speed (m/s)	Barometer (mm Hg)	Barometer Temp. (°C)	Corrected Pressure (hPa)	
Feb 12	7:48	-2.0	-8.5	-7.0	-6.2	F	-6.2	A	--		0	--	BKN	6	1	--	St	As	--	SW	1	239.6	1.962	700.8	20.0		Fog lower, in bowl (not in sky view) Frost on bench
Feb 12	12:30	-2.0	-8.0	-2.0	-3.1	F	-2.1	A	0		0	--	OVC	0	0	8	--	--	Cs	E	2	101	0.4	700.8	21.0		Clouds Changed
Feb 12	3:30	0.0	-3.0	-1.0	-1.5	F	-1.4	A	--		--	--	SCT	--	2	2	--	As	Cl	SW	1	248	1.4	700.4	22		
Feb 13	4:05	-0.5	-13.0	-12.5	-10.4	F	-10.2	A	--		0	--	CLR	--	--	--	--	--	--	SW	0	229	0.4	703.4	21		Frost on bench, Stratus clouds on horizon
Feb 13	12:30	-8.0	-13.0	-8.0	-8.0	F	-7.8	A	0		0	--	BKN	6	--	--	St	--	--	NE	2	81	1.6	703.5	21		Low clouds are changing
Feb 13	15:50	-6.0	-8.0	-6.0	-7.6	F	-6.4	A	--		--	--	Few	1	--	--	Sc	--	--	NE	1	348	0.4	702.6	21		
Feb 13	8:04	-8.0	-15.0	-13.0	-10.9	F	-10.8	A	--		0	--	OBSCD	--	--	--	--	--	--	N	1	16	1.9	702.8	21		Frost on bench, Shelter, railings etc, Fog!
Feb-14	12:21	-11.0	-13.0	-11.0	-9.5	F	-9.5	A	--		0	--	OBSCD	--	--	--	--	--	--	NE	2	16	0.6	702.5	21.0		Fog; can't see anything
Feb 14	3:45	-5.0	-11.0	-5.0	-5.8	F	-5.3	A	--		--	--	CLR	--	--	--	--	--	--	NE	1	348	1.4	702.6	21.0		Frosty
Feb 15	7:35	-4.5	-13.5	-13.0	-11.4	F	-10.6	A	--		0	--	CLR	--	--	--	--	--	--	N	2	358	2.5	703.1	21.0		Frost on bench Fog & stratus clouds in bowl
Feb 15	12:36	-6.0	-13.0	-6.0	-7.2	F	-5.1	A	0		0	--	CLR	0	0	0	--	--	--	N	2	62	0.6	704.2	21.0		Clear sky & variable low speed wind.
Feb 15	3:45	-2.0	-6.0	-2.0	-4.6	F	-2.2	A	0		0	--	CLR	0	0	0	--	--	--	NE	1	52	0.9	705	21.0		Lovely day in

Each student calculates their own e (hPa) & RH (%) values from recorded observations.

Each student calculates their own SWE values from recorded observations.

Each student calculates their own pressures from recorded observations.

DO NOT REMOVE – LEAVE DATA SHEET WITH CLIPBOARD AT ALL TIMES



2024 Today's Lab & What's Coming:

6 / Feb 5	Middle-latitude Cyclones. Weather maps.	Lab 5: 1st turn-in Lab 4: 2nd turn-in Lab 3: returned ***Wx Proj: Barometer Measurements, Calculations & Data Collection Practice Run – Outside…***
7/ Feb 12	Atmospheric stability & cloud formation. Air pollution.	**Lab 6:** Weather Maps & Analysis **(2%)** Lab 6: 1st turn-in Lab 5: 2nd turn-in Lab 4: returned ***WxProj: Interpreting Weather Maps;***
	Wx Proj: Data Collection *Mon to Thu:* ***(6%).*** *Remember your observation time, partner meeting place, key returns.* *Complete:* ***Roof-top Observations*** *(2%);* ***Weather Diary*** *(2%);* ***Electronic Synoptic Data Collection*** *(1%);* ***Teamwork Evaluation*** *(1%)* **Due by 10 am Friday Feb 16 in your dropbox**: Personal weather diary, completed teamwork evaluation, & collected electronic data on a memory key. All are submitted in a properly labelled, sealed ziplock bag.	
Feb 19	**Family Day** (Mon) **& Mid-Semester Break** – no classes Feb 19 – Feb 23	
8 / Feb 26	Condensation, cloud & precipitation formation. ***Wx Proj: How to write a scientific report.*** *Collected Wx Proj data returned in Labs or here*	***WxProj: Data Quality Control,*** **produce Appendix 2, (1%)** **Bring: your laptop or use lab computers.** Appendix 2 submitted Lab 6: 2nd turn-in Lab 5, Collected WxProj data: returned
9 / Mar 4	Forces in the atmosphere. Atmospheric dynamics & wind. Jet streams, upper-level patterns.	***WxProj: Time Series Graphing,*** **produce your report graphs (1%)** ***Sign-up for Report outline meeting times*** Lab 6: returned Appendix 2 returned
10 / Mar 11	Atmosphere / Greenhouse effect. Global climate & climate change.	***WxProj: Report Outline Interview Meetings*** **(2%)** Bring prepared WxProj outline & resources for discussion
11 / Mar 18	Stratospheric ozone.Tropical cyclones.	**No labs – work on Wx Project**
12 / Mar 25	Severe summer weather.	**No labs – work on Wx Project**
13 / Apr 1	Severe winter weather. Exam prep & course review.	**No labs**
	WxProj Report Due: Tue Apr 2 at 4 pm in your drop box & electronically on Moodle (20%). **Late reports lose 20% per day** (including weekends & holidays) **staring after the due time.**	
Exam Period April 10 - 20, 2024. ENSC 201 In-person Final Exam (33%), date TBD.		

Today's Agenda:

Returned: **Lab 5, Lab 6, WxProj Data Collection Submission** (electronic data, diary & data collection marks). **Return your ziplock bag** for re-use, if you don't need it.

Handouts: Lab 6 Reflection **(blue sheet)**
1 photocopy of your original complete roof-top datasheet
1 blank final datasheet

Required: Read WxProj Report Requirements (WxProj Guide pp. 9 -14) this week.

1) **Appendix 1 & 2 are part of your Wx Proj Report** and must be properly included in your report's paper copy.

2) **Produce Appendix 2 – done in today's lab.** *See next slide for more.*

3) **Produce Appendix 1** (your marked WxProj Diary) **on your own time. Keep it!**
Format your marked weather diary for including in your report as Appendix 1.
   - include it as marked with comments; do not revise /redo, it is not graded again
   - follow a consistent Appendix format in your report (should fit with Appendix 2)
   - See WxProj guide pp. 2, 9, 10; your instructor as needed

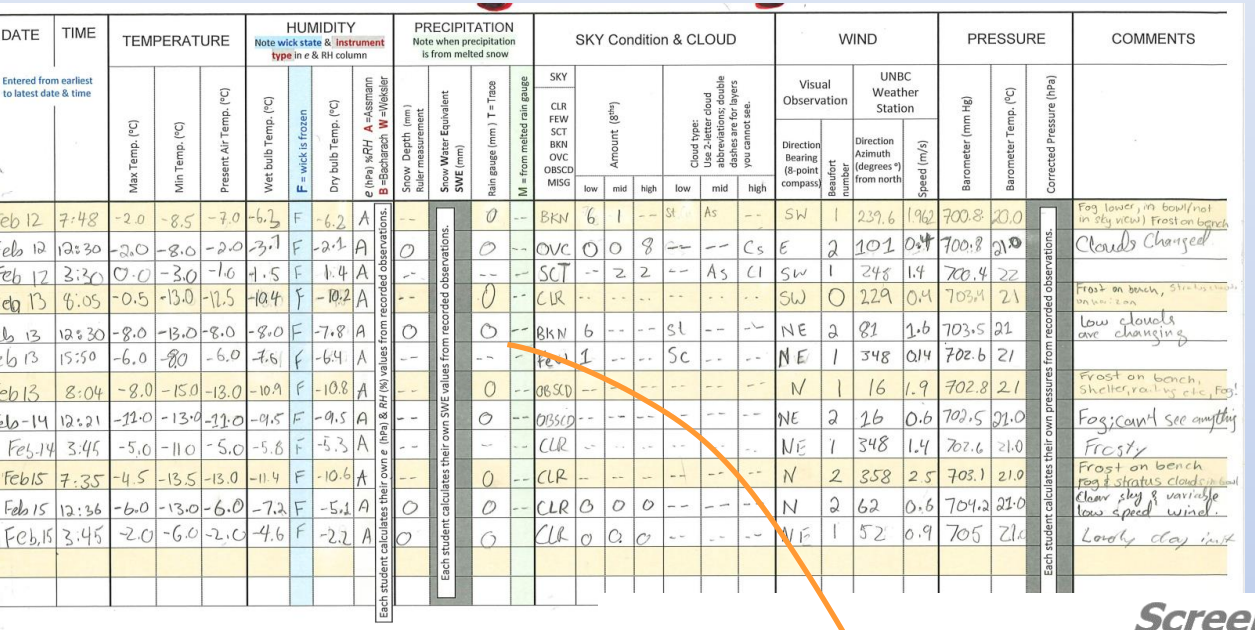
Due ~2 days (regular dropbox deadline): **Lab 6** 2nd submission **& Appendix 2**

Today's Lab → Data Rationalization & creating Appendix 2

What is data rationalization: The process of transforming raw data into a final dataset (involves data quality control & calculating final values); and documenting this process so report readers understand the data & its evolution (for us this is Appendix 2).

- Clearly document & show your data rationalization process. Your documentation should justify your quality control decisions and be detailed enough for readers to follow each data issue and decision. Calculations should be detailed enough to follow unit conversions and show how results are obtained.
- The concept of data rationalization:
  - → Presents your raw data and clearly indicates specific data issues (use colour and footnotes) to indicate which data points are affected
  - → Lists identified data point issues, explains them (their causes /impacts), and indicates what was specifically done to rectify them (if that is possible, but if not possible or worth doing, it indicates nothing was done, the reasoning why, and the consequences of doing nothing).
  - → Calculates computed values from raw measurements. Provides an example calculation for each type.
  - → Presents your final dataset showing which data points have changed if any are altered

Data Rationalization:



Transform raw data into the final data set.

Must ensure it is the highest quality it can be before any further analysis.

Final data has undergone data rationalization
(data quality control & raw values are converted to their final form).

Screen *Instrument*

DATE	TIME	TEMPERATURE			HUMIDITY					PRECIPITATION				SKY Condition & CLOUD								WIND				
Entered from earliest to latest date & time					Note wick state & instrument type in the indicated columns					Note when precipitation is from melted rain gauge																
2024 Data		Max Temp. (°C)	Min Temp. (°C)	Present Air Temp. (°C)	Wet bulb Temp. (°C)	F = wick is frozen	Dry bulb Temp. (°C)	e (hPa)	RH (%) A = Assmann B = Bacharach W = Weksler	Snow Depth (mm) Ruler measurement	Snow Water Equivalent SWE (mm)	Rain gauge (mm) T = Trace	M = from melted rain gauge	SKY CLR FEW SCT BKN OVC OBSCD MISG	Amount (8ths) low	Amount (8ths) mid	Amount (8ths) high	Cloud type: Use 2-letter cloud abbreviations & double dashes for layers you cannot see. low	Cloud type: Use 2-letter cloud abbreviations & double dashes for layers you cannot see. mid	Cloud type: Use 2-letter cloud abbreviations & double dashes for layers you cannot see. high	Visual Observation Direction Bearing (8-point compass)	Visual Observation Beaufort number	UNBC Weather Station Direction Azimuth (degrees °) from north	UNBC Weather Station Speed (m/s)	Barometer (mm Hg)
Mon Feb 12						F			A				--													
Mon Feb 12						F			A				--													
Mon Feb 12						F			A				--													
Tue Feb 13						F			A				--													
Tue Feb 13						F			A				--													
Tue Feb 13						F			A				--													
Wed Feb 14						F			A				--													
Wed Feb 14						F			A				--													
Wed Feb 14						F			A				--													
Thu Feb 15						F			A				--													
Thu Feb 15						F			A				--													

Appendix 2 Required Format & Elements, in this order:

1. Title (or title page) & descriptive statement of purpose (see your textbook /Course Manual for examples)

2. Final rationalized dataset with data issues indicated as footnotes & colour-coded cells that match the data problems identified in your raw (original) data set

3. Your footnoted list of issues, their impacts & their resolution. This is a separate page that tracks your data issues, explains them, indicates their impact on your report, and shows how you dealt with each problem. It may include indicating limitations on the use of specific data.

   Your list of footnotes should show how the data issue in your raw dataset links with its resolution in the final dataset and provide the information to show its evolution from original raw data to final data.

4. Sample calculations, one for each type of calculated value (show how your raw measurement becomes a final dataset value).

5. Raw (original) dataset with data problems indicated as footnotes & colour-coded cells that match the footnoted list of data issues & identified final dataset cells

Remember: Appendices have page numbers…e.g. 2-1, 2-2, etc.

What exactly is an error?

Objective vs Subjective Errors?

Data rationalization – types of errors

Objective errors - data impossibilities

- Max-Min-Present air temperature correspondence.
  - Max less than (<) Min;
  - Min greater than (>) Max;
  - Present air temperatures not equal to or between max & min
- Wet bulb temperatures > dry bulb temperatures
- Sky condition & cloud amounts don't agree
- Missing data? depends on why is it not there…..

Subjective errors – data you suspect but can't prove wrong

- Suspect data: Odd values /data inconsistencies but you don't have objective evidence to show /verify the value is incorrect, e.g. temperatures that seem too low /too high, or don't follow anticipated patterns
- Cloud types that seem unlikely given the sky /weather conditions, diary
- Odd discrepancies between anemometer & visual wind observations
- Others?

Data rationalization – Identify & track

- **Clearly show how you obtain your final dataset!**

  Use highlighting to draw attention to affected cells on both raw and final datasets.

  Colours can make footnotes easier to see!

- Explicitly identify, annotate cells with a footnote number,
- Create a list & explain each data issue or type of issue. Report how data issues were treated /resolved.
- **Use:**
  - the same footnotes (1, 2, 3) in cells with issues on both the raw & final data tables
  - list each footnote & explain the data issue, & its resolution

  → sub-notes (1a, 1b) can be useful to indicate multiple errors of the same type

What should you do with problem data?

Replace?

Leave as is?

Data rationalization – Solutions cont.

Identify appropriate solutions:

1. **Understand your data errors** - it may help to speak with those who collected your data /instructors

2. Seek **objective** rather than **subjective** solutions

3. Consider your **solution's impact on your use of the data**

4. Be transparent: **Track & clearly explain your process** (the error, its resolution (if any), consequences of your resolution, your decision)

Subjectively changing data is never appropriate!

An undocumented memory of your measurement is not objective & is a poor solution for problem data (it is not transparent enough).

Replace?

As we have data replication, we can possibly recover data by replacing it (though not always). Consider the consequences in each case.

When replacing objectively incorrect data with a best-matched value from a replicate dataset, this MUST be documented & replacement(s) must improve the data issue!

- Look for best matches - time & instrument compatibility (readings)
- Explicitly identify each data issue - state the problem (explain the error whenever possible), show which data point(s) are affected
- Track & explain your solution - state your resolution (report the source of each replacement data value, use footnotes to link this information to the original raw data and your list of data problems)

Problem Data Solutions

Sometimes there are no good solutions….

Leave as is?

Yes, but:

- data **problems must still be identified & tracked,** sometimes they can limit interpretation affecting your understanding of the weather story
- still must **understand & consider the limitations of your data** when interpreting & analyzing it.

Seek instructor assistance with these issues

Data rationalization – Calculating final values

Compute final values & Example Calculations

(as done in labs)

- ***e*** **& Relative Humidity** (track your e*, T, or T_w values)
- **Precipitation** (convert snowfall measurements to water equivalents)
- **Pressure** (report as a corrected station pressure in hPa)

Use Excel to speed calculating /track values. Then record computed data values on your final datasheet in the appropriate columns.

Create Appendix 2 (required report component)

- Put required Appendix 2 elements in the right order
- Add appropriate titles, statements of purpose, page numbers
- Revisit Appendix 2 requirements slides so nothing is missed…..

So, what is an Appendix?

- See your textbook or Course Manual for examples
- Appendices supplement a report without interrupting the report's focus or flow.
- Like figures, each appendix must be descriptively titled & have a statement of purpose (like a caption) to make the purpose / use of the appendix clear to the reader
- When appendices have multiple components, each component must have a title & caption
- Like figures, every appendix must be referred to in the body of your report
- Appendices have page numbers (e.g. carry on from report, or 1-1, 1-2; 2-1, 2-2, etc.)

Recap: What is Appendix 2?

- Presents the final and original raw dataset used in your report
- Transparently shows the evolution of your original raw data (it's quality control measures and calculations) to its final form.
- Lists your specific data issues, their impact, & their resolution
  Presents your final & raw (original) data, identifies any data issues, and shows /explains data problem resolutions. Appendix 2 makes it possible to understand the evolution /progression of your data.
- Shows data calculation details

Appendices 1 & 2 are required report elements.

- Appendix 1: your returned & graded weather diary in its original form
- Appendix 2: your rationalized dataset and its evolution (from raw to final with clear documentation). Follows the required format.
- Hand written

Recap: Appendix 2 Elements (in this order)

- **Title & statement of purpose** e.g. "Appendix 2…. & a descriptive title", a brief description of Appendix 2's purpose.
- **Final, rationalized dataset** with footnotes /symbols that match those in your list of errors, & raw dataset clearly showing your data's evolution.
- **A list of data issues & errors**, their impacts /explanation, & your tracking of their resolution /how you will deal with each data problem. This list should match your dataset footnotes /symbols and link your list to the data issues.
- **Sample calculations** showing how you computed final values
- **Original raw rooftop dataset** with footnotes or symbols that match those in your list of errors, & final dataset.

Remember to: Indicate linked changes; use the same footnotes /symbols /colour-coding on your original dataset, listing of errors, & final dataset. A legend may be useful for symbols. Give your appendices page numbers.

This week:

Due in your dropbox by your lab assignment deadline:

- **Appendix 2** & **Lab 6** (2nd submission)

Next week:

1) Stacked time series graphing (worth 1%; marked in lab)

   Must show your draft graph by the end of the lab & bring it for use in your instructor interview.

2) Set your Wx Proj Instructor meeting times for the following week!

Presentation END.