St3TART
FRM for Sentinel-3 Land Altimetry

Sentinel-3 Topography mission Assessment through Reference Techniques (St3TART)

TD-3 FRM Protocols and Procedure for S3 STM Land Ice Products

	Function	Name	Signature	Date
Prepared by	St3TART Project Team	Geir MOHOLDT		17/04/23
Approved by	Project manager	Elodie Da Silva		17/04/23
Authorized by	Deputy CEO	Mahmoud EL HAJJ		17/04/23
Accepted by	ESA Technical Officer	Pierre FEMENIAS		

NOVELTIS

afaq
ISO 9001
Qualité
AFNOR CERTIFICATION

Distribution list

INTERNAL	EXTERNAL	
Name	**Name**	**Company / Organisation**
NOVELTIS Documentation	Mr. Pierre FÉMÉNIAS	ESA
Mr. Richard BRU	Mr. Jérôme BOUFFARD	ESA
Mr. Mahmoud EL HAJJ	Mr. Philippe GORYL	ESA
St3TART Project team	Ms Filomena CATAPANO	RHEA for ESA

Document status

Sentinel-3 Topography mission Assessment through Reference Techniques (St3TART) Deliverable Name			
Issue	**Revision**	**Date**	**Reason for the revision**
1	0	27/10/2021	First version
2	0	05/05/2022	Second version
3	0	13/01/2023	Third version
4	0	03/04/2023	Final version
4	1	17/04/2023	Updates following ESA RIDs

Modification status				
Issue	Rev	Status *	Modified pages	Reason for the modification
2	0	M	All	Second version
3	0	M	All	Third version
4	0	M	All	Final version
4	1	M	Almost all	Minor changes all over the document to take into account ESA RIDs

*   I = Inserted   D = Deleted   M = Modified

Acronyms

AD	Applicable Document
AIPo	Agenzia Interregionale del Fiume Po
AOI	Area Of Interest
ALS	Airborne Laser Scanner
ATM	Airborne Topographic Mapper
AUV	Autonomous Underwater Vehicle
BVLoS	Beyond Visible Line of Sight
CEN	Centre d'Etudes de la Neige
CGLS	Copernicus Global Land Service
CO	ESA Contract Officer
CS-2	CryoSat-2 mission
CSV	Comma-Separated Values
DSM	Digital Surface Model (Digital Elevation Model)
DMS	Digital Mapping System
EEA	European Environment Agency
EM	Airborne Elextromagnetic Induction
EOB	End Of Business
ESA	European Space Agency
EU	European Union
FB	Sea Ice Freeboard
FFP	Firm Fixed Price
FFSAR	Fully Focused SAR
FIR	Flight Information Regions
FM-CW	Frequency-modulated continuous-wave radar
FOEN	Federal Office for the Environment
FR	Final Review
FRM	Fiducial Reference Measurement
GCP	Ground Control Points
GEM	Ground based ElectroMagnetic
GNSS	Global Navigation Satellite System
GPR	Ground Penetrating Radar
GPS	Global Positioning System
GRDC	Global Runoff Data Centre

GRP	Ground Reference Point
HDD	Hard Disk Drive
HR	High Resolution
IMB	Ice Mass Balance
INS	Inertial Navigation System
IOCR	In-Orbit Commissioning Review
ISPRA	Instituto Superiore per la Protezione e la Ricerca Ambientale
KO	Kick-Off
KOM	Kick-Off Meeting
L1	Level 1
L2	Level 2
LAM	Low Altitude Mode
LOCSS	Lake Observations by Citizen Scientists & Satellites (OECS in France)
LVIS	Land, Vegetation and Ice Sensor
LRM	Low Resolution Mode
MFF	Multi-annual Financial Framework
MoM	Minutes of Meeting
MPC	S3 Mission Performance Cluster
MSS	Mean Sea Surface
MVP	Minimum Viable Product
N/A	Not Applicable
NMI	National Metrology Institute
NWD	Normal Working Day
OC	Operational Challenges
OCOG	Offset Centre Of Gravity
OIB	Nasa Operation IceBridge
OLTC	Open Loop Tracking Commands table
OS	Operating System
OSSE	Observing System Simulation Experiments
PACF	Permanent Altimeter Calibration Facility
PDF	Portable Document Format
PM	Progress Meeting
PMP	Project Management Plan
POCA	Point Of Closest Approach
PPK	Post Processed Kinematic
PPP	Precise Point Positionning

PRF	Pulse Repetion Frequency
QA4EO	Quality Assurance Framework for Earth Observation
QWG	Quality Working Group
R&D	Research & Development
RD	Reference Document
RO	Routine Operational Phase
RTK	Real-Time Kinematic positioning
RULS	Range Under the Level of the Sonic gauge
S3	Copernicus Sentinel-3 mission
S3VT	Sentinel-3 Validation Team
S4B	Skype For Business
S6	Copernicus Sentinel-6 mission
SAMS	Scottish Association for Marine Sciences
SAR	Synthetic Aperture Radar
SCalSIT	Super CalVal Site Identifier Tool
SI	Système International
SILDAMS	Sea-Ice Lead Detection Algorithm using Minimal Signal
SIMBA	Snow and Ice Mass Balance Array
SLA	Sea Level Anomaly
SMP	SnowMicroPen
SRAL	Synthetic Aperture Radar Altimeter
SIRAL	CryoSat's SAR Interferometric Radar Altimeter (SIRAL)
SSH	Sea Surface Height
St3TART	Sentinel-3 Topography mission Assessment through Reference Techniques
STM	Surface Topography Mission
SWOT	Surface Water and Ocean Topography mission
TC	Technological Challenges
TD	Technical Deliverable
TO	ESA Technical Officer
UAV	Unmanned Aerial Vehicle
ULS	Upward Looking Sonar
UWB	Ultra-Wideband
VC	Video-Conference
WP	Work Package

Reference documents

N°	Reference	Title
[RD1]	ESA-EOPG-CSCOP-SOW-29, Issue 1 Rev. 4 – 28/10/2020	Statement of Work - Sentinel-3 Topography mission Assessment through Reference Techniques (St3TART)
[RD2]	NOV-FE-0899-PR-002	Technical Proposal
[RD3]	NOV-FE-0899-PR-004	Implementation Proposal
[RD4]	ESA Contract No. 4000135181/21/I-DT	ESA Contract – Copernicus ground segment Sentinel-3 Topography mission Assessment through Reference Techniques (St3TART)
[RD5]	NOV-FE-0899-NT-040	Acronyms list
[RD6]	https://sentinel.esa.int/nl/web/sentinel/user-guides/sentinel-3-altimetry	Sentinel-3 Altimetry USer Guide
[RD7]	https://sentinel.esa.int/nl/web/sentinel/technical-guides/sentinel-3-altimetry/	Sentinel-3 Altimetry Technical Guide
[RD8]	S3MPC-STM_RP_0038 (Issue 1.1)	Sentinel-3 SRAL/MWR Land User Handbook
[RD9]	CLS-ENV-NT-22-035	St3TART WP 3.2.3: Cross-Platform validation with ICESat-2
[RD10]	CLS-ENV-NT-23-0001	St3TART WP 3.2.4: HR-DEM applications for land ice altimetry
[RD11]	Adams, R.J., Brown, G.S., 1998: A model for altimeter returns from penetrable geophysical media. IEEE Trans. Geosci. Remote Sens. 36, 1784–1793. https://doi. org/10.1109/36.718645.	
[RD12]	Adams, M. S., Bühler, Y., and Fromm, R., 2018: Multitemporal Accuracy and Precision Assessment of Unmanned Aerial System Photogrammetry for Slope-Scale Snow Depth Maps in Alpine Terrain, Pure Appl. Geophys., 175, 3303–3324, https://doi.org/10.1007/s00024-017-1748-y.	
[RD13]	Adodo, F. I., F. Remy, and G. Picard, 2018: Seasonal variations of the backscattering coefficient measured by radar altimeters over the Antarctic Ice Sheet. The Cryosphere, 12, 1767–1778.	
[RD14]	Agosta, C., V. Favier, C. Genthon, H. Gallée, G. Krinner, J. T. M. Lenaerts, and M. R. van den Broeke, 2012: A 40-year accumulation dataset for Adelie Land, Antarctica and its application for model validation. Clim. Dyn., 38, 75–86.	
[RD15]	Agosta, C., and Coauthors, 2019: Estimation of the Antarctic surface mass balance using the regional climate model MAR (1979–2015) and identification of dominant processes. The Cryosphere, 13, 281–296.	
[RD16]	Amory, C., 2020: Drifting-snow statistics from multiple-year autonomous measurements in Adélie Land, East Antarctica, The Cryosphere, 14, 1713–1725.	
[RD17]	Amory, C., F. Naaim-Bouvet, H. Gallée, and E. Vignon, 2016: Brief communication: Two well-marked cases of aerodynamic adjustment of sastrugi. The Cryosphere, 10, 743–750.	
[RD18]	Amory, C., and Coauthors, 2017: Seasonal Variations in Drag Coefficient over a Sastrugi-Covered Snowfield in Coastal East Antarctica. Bound.-Layer Meteorol.	

[RD19]	Amory, C., Kittel, C., L. Le Toumelin, C. Agosta, A. Delhasse, V. Favier, and X. Fettweis, 2021: Performance of MAR (v3.11) in simulating the drifting-snow climate and surface mass balance of Adélie Land, East Antarctica. Geosci. Model Dev., 14, 3487–3510.
[RD20]	Anschütz, H., Müller, K., Isaksson, E., McConnell, J. R., Fischer, H., Miller, H., Albert, M., and Winther, J. G.: Revisiting sites of the South Pole Queen Maud Land Traverses in East Antarctica: Accumulation data from shallow firn cores, J. Geophys. Res., 114, D24106, doi:10.1029/2009JD012204, 2009
[RD21]	Arcone, S.A. (1996), High resolution of glacial ice stratigraphy: A ground-penetrating radar study of Pegasus Runway, McMurdo Station, Antarctica. Geophysics. 61(6): 1653-1663
[RD22]	Arnaud, L., G. Picard, N. Champollion, F. Domine, J. C. Gallet, E. Lefebvre, M. Fily, and J. M. Barnola, 2011: Measurement of vertical profiles of snow specific surface area with a 1 cm resolution using infrared reflectance: instrument description and validation. J. Glaciol., 57, 17–29.
[RD23]	Avanzi, F., Bianchi, A., Cina, A., De Michele, C., Maschio, P., Pagliari, D., Passoni, D., Pinto, L., Piras, M., and Rossi, L., 2018: Centimetric Accuracy in Snow Depth Using Unmanned Aerial System Photogrammetry and a MultiStation, Remote Sensing, 10, 765, https://doi.org/10.3390/rs10050765.
[RD24]	Bamber, J. and Gomez-Dans, J. L., 2005: The accuracy of digital elevation models of the Antarctic continent, Earth Planet. Sc. Lett., 237, 3–4, 516–523.
[RD25]	Bamber, J. L., J. L. Gomez-Dans, and J. A. Griggs, 2009: A new 1 km digital elevation model of the Antarctic derived from combined satellite radar and laser data – Part 1: Data and methods. Cryosphere Discuss, 2, 811–841.
[RD26]	Bindschadler, R.A. and Vornberger, P.L.. 1994. Detailed elevation map of Ice Stream C, Antarctica, using satellite imagery and airborne radar. Ann. Glaciol., 20, 327–335
[RD27]	BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML (2008), Evaluation of measurement data—Guide to the expression of uncertainty in measurement, International Organization for Standardization (ISO), Online: http://www.bipm.org/en/publications/guides/gum.html.
[RD28]	Blair, J. B. and Hofton, M., 2010: IceBridge LVISL2 Geologcated Surface Elevation Product, Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center, updated 2012.
[RD29]	Brenner, A.C., Bindschadler, R.A., Thomas, R.H., Zwally, H.J., 1983: Slope-induced errors in radar altimetry over continental ice sheets. J. Geophys. Res. 88, 1617–1623. https://doi.org/10.1029/JC088iC03p01617.
[RD30]	Brown, G.S., 1977: The average impulse response of a rough surface and its applications. IEEE J. Ocean. Eng. 2, 67–74. https://doi.org/10.1109/JOE.1977.1145328.
[RD31]	Bühler, Y., Adams, M. S., Bösch, R., and Stoffel, A., 2016: Mapping snow depth in alpine terrain with unmanned aerial systems (UASs): potential and limitations, The Cryosphere, 10, 1075–1088, https://doi.org/10.5194/tc-10-1075-2016.
[RD32]	Cohen, L., and S. Dean, 2013: Snow on the Ross Ice Shelf: comparison of reanalyses and observations from automatic weather stations. The Cryosphere, 7, 1399–1410.
[RD33]	Crocker, R.I., J. A. Maslanik, J. J. Adler, S. E. Palo, U. C. Herzfeld and W. J. Emery, "A sensor package for ice surface observations using small unmanned aircraft systems", *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1033-1047, Apr. 2012.
[RD34]	Deems, J. S., Painter, T. H., and Finnegan, D. C., 2013: Lidar measurement of snow depth: a review, J. Glaciol., 59, 467–479.
[RD35]	De Michele, C., Avanzi, F., Passoni, D., Barzaghi, R., Pinto, L., Dosso, P., Ghezzi, A., Gianatti, R., and Della Vedova, G., 2016: Using a fixed-wing UAS to map snow depth distribution: an evaluation at peak accumulation, The Cryosphere, 10, 511–522, https://doi.org/10.5194/tc-10-511-2016.
[RD36]	Ding, M., C. Xiao, Y. Li, J. Ren, S. Hou, B. Jin, and B. Sun, 2011: Spatial variability of surface mass balance along a traverse route from Zhongshan station to Dome A, Antarctica. J. Glaciol., 57, 658–666.

[RD37]	Ducasse, E., E. Berthier, D. Blumstein, E. Le Meur, F. Gillet-Chaulet, G. Durand, "Recent elevation and velocity changes of Astrolabe Glacier, Terre Adelie, Antarctica," 2015 8th International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi-Temp), Annecy, France, 2015, pp. 1-4, doi: 10.1109/Multi-Temp.2015.7245769.
[RD38]	Edwards, T. L., and Coauthors, 2019: Revisiting Antarctic ice loss due to marine ice-cliff instability. Nature, 566, 58, https://doi.org/10.1038/s41586-019-0901-4.
[RD39]	Eisen, O., Nixdorf, U., Wilhelms, F., and Miller, H.: Age estimates of isochronous reflection horizons by combining ice core, survey, and synthetic radar data, J. Geophys. Res., 109, B04106, https://doi.org/10.1029/2003JB002858, 2004.
[RD40]	Eisen, O., and Coauthors, 2008: Ground-based measurements of spatial and temporal variability of snow accumulation in East Antarctica. Rev. Geophys., 46, RG2001.
[RD41]	Favier, V., C. Agosta, C. Genthon, L. Arnaud, A. Trouvillez, and H. Gallée, 2011: Modeling the mass and surface heat budgets in a coastal blue ice area of Adelie Land, Antarctica. J. Geophys. Res. Earth Surf., 116, F03017.
[RD42]	Favier V., and Coauthors, 2013: An updated and quality controlled surface mass balance dataset for Antarctica. The Cryosphere, 7, 583–597.
[RD43]	Femenias, P., Remy, F., Raizonville, R., Minster, J.F., 1993: Analysis of satellite-altimeter height measurements above continental ice sheets. J. Glaciol. 39, 591–600. https://doi.org/10.3189/S0022143000016488.
[RD44]	Filhol, S. and Sturm, M., 2015: Snow bedforms: A review, new data and a formation model, J. Geophys. Res., 120, 1645–1669, doi:10.1002/2015JF003529.
[RD45]	Fountain, A. G., T. H. Nylen, A. Monaghan, H. J. Basagic, and D. Bromwich, 2010: Snow in the McMurdo Dry Valleys, Antarctica. Int. J. Climatol., 30, 633–642.
[RD46]	Fretwell, and Coauthors, 2013: Bedmap2: improved ice bed, surface andthickness datasets for Antarctica, The Cryosphere, 7, 375–393.
[RD47]	Fréville, H., E. Brun, G. Picard, N. Tatarinova, L. Arnaud, C. Lanconelli, C. Reijmer, and M. van den Broeke, 2014: Using MODIS land surface temperatures and the Crocus snow model to understand the warm bias of ERA-Interim reanalyses at the surface in Antarctica. The Cryosphere, 8, 1361–1373.
[RD48]	Frezzotti, M., Pourchet, M., Flora, O., Gandolfi, S., Gay, M., Urbini, S., Vincent, C., Becagli, S., Gragnani, R., and Proposito, M.: Spatial and temporal variability of snow accumulation in East Antarctica from traverse data, J. Glaciol., 51, 113–124, 2005
[RD49]	Gallée, H., G. Guyomarc'h, and E. Brun, 2001: Impact Of Snow Drift On The Antarctic Ice Sheet Surface Mass Balance: Possible Sensitivity To Snow-Surface Properties. Bound.-Layer Meteorol., 99, 1–19.
[RD50]	Gallée, H., A. Trouvilliez, C. Agosta, C. Genthon, V. Favier, and F. Naaim-Bouvet, 2013: Transport of Snow by the Wind: A Comparison Between Observations in Adélie Land, Antarctica, and Simulations Made with the Regional Climate Model MAR. Bound.-Layer Meteorol., 146, 133–147.
[RD51]	Genthon, C., L. Piard, E. Vignon, J.-B. Madeleine, M. Casado, and H. Gallée, 2017: Atmospheric moisture supersaturation in the near-surface atmosphere at Dome C, Antarctic Plateau. Atmospheric Chem. Phys., 17, 691–704, https://doi.org/10.5194/acp-17-691-2017.
[RD52]	Gindraux, S., Boesch, R., and Farinotti, D., 2017: Accuracy Assessment of Digital Surface Models from Unmanned Aerial Vehicles' Imagery on Glaciers, Remote Sensing, 9, 186, https://doi.org/10.3390/rs9020186.
[RD53]	Goetz, J. and Brenning, A., 2019: Quantifying Uncertainties in Snow Depth Mapping From Structure From Motion Photogrammetry in an Alpine Area, Water Resources Research, 55, 7772–7783, https://doi.org/10.1029/2019WR025251.

[RD54]	Griggs, J. A. and Bamber, J. L., 2009: A new 1 km digital elevation model of Antarctica derived from combined radar and laser data – Part 2: Validation and error estimates, The Cryosphere, 3, 113–123.
[RD55]	Guerreiro, K., Fleury, S., Zakharova, E., Remy, F., Kouraev, A., 2016: Potential for estimation of snow depth on Arctic sea ice from CryoSat-2 and SARAL/AltiKa missions. Remote Sens. Environ. 186, 339–349. https://doi.org/10.1016/j. rse.2016.07.013.
[RD56]	Guerreiro, K., Fleury, S., Zakharova, E., Kouraev, A., Remy, F., Maisongrande, P., 2017: Comparison of cryosat-2 and envisat radar freeboard over arctic sea ice: toward an improved envisat freeboard retrieval. Cryosphere 11, 2059–2073. https://doi.org/ 10.5194/tc-11-2059-2017.
[RD57]	Harder, P., Pomeroy, J. W., and Helgason, W. D., 2020: Improving sub-canopy snow depth mapping with unmanned aerial vehicles: lidar versus structure-from-motion techniques, The Cryosphere, 14, 1919–1935, https://doi.org/10.5194/tc-14-1919-2020.
[RD58]	Hawley, R. L., E. M. Morris, R. Cullen, U. Nixdorf, A. P. Shepherd, D. J. Wingham, "ASIRAS airborne radar resolves internal annual layers in the dry-snow zone of Greenland," Geophysical Research Letters, Vol. 33, L04502, doi:10.1029/2005GL025147, 2006
[RD59]	Hawley, R., Brandt, O., Dunse, T., Hagen, J., Helm, V., Kohler, J., . . . Høgda, K. (2013). Using airborne Ku-band altimeter waveforms to investigate winter accumulation and glacier facies on Austfonna, Svalbard. Journal of Glaciology, 59(217), 893-899. doi:10.3189/2013JoG13J051
[RD60]	Helm, V., Humbert, A., and Miller, H., 2014: Elevation and elevation change of Greenland and Antarctica derived from CryoSat-2, The Cryosphere, 8, 1539–1559.
[RD61]	Hofton, M. A., Blair, J. B., Luthcke, S. B., and Rabine, D. L., 2008: Assessing the performance of 20–25 m footprint waveform lidar data collected in ICESat data corridors in Greenland, Geophys.Res. Lett., 35(24).
[RD62]	Howat, I.M., Porter, C., Smith, B.E., Noh, M.J., Morin, P., 2019. The reference elevation model of Antarctica. Cryosphere 13, 665–674.
[RD63]	IMBIE team, 2018: Mass balance of the Antarctic Ice Sheet from 1992 to 2017. Nature, 558, 219–222.
[RD64]	Knuth, S.L., 2007: Estimation of Snow Accumulation in Antarctica Using Automated Acoustic Depth Gauge Measurements. M.Sc. Thesis, University of Wisconsin-Madison, Madison.
[RD65]	Konrad, H., Hogg, A. E., Mulvaney, R., Arthern, R., Tuckwell, R. J., Medley, B., and Shepherd, A.: Observations of surface mass balance on Pine Island Glacier, West Antarctica, and the effect of strain history in fast-flowing sections, J. Glaciol., 65, 1–10, https://doi.org/10.1017/jog.2019.36, 2019
[RD66]	Kowalewski, S., Helm, V., Morris, E. M., and Eisen, O.: The regional-scale surface mass balance of Pine Island Glacier, West Antarctica, over the period 2005–2014, derived from airborne radar soundings and neutron probe measurements, The Cryosphere, 15, 1285–1305, https://doi.org/10.5194/tc-15-1285-2021, 2021
[RD67]	Krabill, W. B., Thomas, R. H., Martin, C. F., Swift, R. N., & Frederick, E. B. (1995). Accuracy of airborne laser altimetry over the Greenland ice sheet. International Journal of Remote Sensing, 16(7), 1211–1222. https://doi.org/10.1080/01431169508954472
[RD68]	Krabill, W. B., 2010: IceBridge ATML1B Qfit Elevation and Return Strength, Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center, updated 2013.
[RD69]	Kuipers Munneke, P., McGrath, D., Medley, B., Luckman, A., Bevan, S., Kulessa, B., et al. (2017). Observationally constrained surface mass balance of Larsen C ice shelf, Antarctica. The Cryosphere, 11, 2411–2426. https://doi.org/10.5194/tc-11-2411-2017
[RD70]	Kurtz, N.T., Galin, N., Studinger, M., 2014: An improved CryoSat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting. Cryosphere 8, 1217–1237. https://doi.org/10.5194/tc-8-1217-2014.

[RD71]	Lampkin, D.J., Vanderberg, J., 2011: A preliminary investigation of the influence of basal and surface topography on supraglacial lake distribution near Jakobshavn Isbrae, western Greenland. Hydrological Processes 25, 3347–3355.
[RD72]	Larue, F., and Coauthors, 2021: Radar altimeter waveform simulations in Antarctica with the Snow Microwave Radiative Transfer Model (SMRT). Remote Sens. Environ., 263, 112534.
[RD73]	Lazzara, M. A., G. A. Weidner, L. M. Keller, J. E. Thom, and J. J. Cassano, 2012: Antarctic Automatic Weather Station Program: 30 Years of Polar Observation. Bull. Am. Meteorol. Soc., 93, 1519–1537.
[RD74]	Legrésy, B., Rémy, F., 1997: Altimetric observations of surface characteristics of the antarctic ice sheet. J. Glaciol. 43, 265–275. https://doi.org/10.3189/ S002214300000321X.
[RD75]	Leuschen, C, Gogineni, S and Tammana, D (2000) SAR processing of radar echo sounder data. In Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International, volume 6, 2570–2572, IEEE
[RD76]	Le Meur, E., and Coauthors, 2014: Two independent methods for mapping the grounding line of an outlet glacier – an example from the Astrolabe Glacier, Terre Adélie, Antarctica. The Cryosphere, 8, 1331–1346.
[RD77]	Le Meur, E., Magand, O., Arnaud, L., Fily, M., Frezzotti, M., Cavitte, M., Mulvaney, R., and Urbini, S.: Spatial and temporal distributions of surface mass balance between Concordia and Vostok stations, Antarctica, from combined radar and ice core data: first results and detailed error analysis, The Cryosphere, 12, 1831–1850, https://doi.org/10.5194/tc-12-1831-2018, 2018.
[RD78]	Li, L., and J. W. Pomeroy, 1997: Estimates of threshold wind speeds for snow transport using meteorological data. J. Appl. Meteorol., 36, 205–213.
[RD79]	Libois, Q., G. Picard, L. Arnaud, S. Morin, and E. Brun, 2014: Modeling the impact of snow drift on the decameter-scale variability of snow properties on the Antarctic Plateau. J. Geophys. Res. Atmospheres, 119, 11,662–11,681.
[RD80]	Ligtenberg, S. R. M., M. M. Helsen, and M. R. van den Broeke, 2011: An improved semi-empirical model for the densification of Antarctic firn. The Cryosphere, 5, 809–819.
[RD81]	MacGregor, J. A., Boisvert, L. N., Medley, B., Petty, A. A., Harbeck, J. P., Bell, R. E., et al. (2021). The scientific legacy of NASA's Operation IceBridge. Reviews of Geophysics, 59, e2020RG000712. https://doi. org/10.1029/2020RG000712
[RD82]	Mann, G. W., P. S. Anderson, S. D. Mobbs, 2000: Profile measurements of blowing snow at Halley, Antarctica, J. Geo- phys. Res., 105(D19), 24,491–24,508.
[RD83]	Martin, C. F., Krabill, W. B., Manizade, S., Russell, R., Sonntag, J. G., Swift, R. N., and Yungel, J. K., 2012: Airborne Topographic Mapper Calibration Procedures and Accuracy Assessment, Rep. 20120008479, Natl. Aeronaut. and Space Admin., Washington, D.C.
[RD84]	Mavrocordatos, C., E. Attema, M. Davidson, H. Lentz and U. Nixdorf, "Development of ASIRAS (Airborne SAR/Interferometric Altimeter System)," IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, USA, 2004, pp. 2465-2467 vol.4, doi: 10.1109/IGARSS.2004.1369792.
[RD85]	Mayewski, P. A., Frezzotti, M., Bertler, N., Van Ommen, T., Hamilton, G., Jacka, T. H., Welch, B., Frey, M., Dahe, Q., and Jiawen, R.: The international trans-antarctic scientific expedition (ITASE): an overview, Ann. Glaciol., 41, 180–185, 2005.
[RD86]	McMillan, M., et al., 2016: A high-reso-lution record of Greenland mass balance,Geophys. Res. Lett.,43, 7002–7010,doi:10.1002/2016GL069666
[RD87]	Medley, B., Joughin, I., Smith, B. E., Das, S. B., Steig, E. J., Conway, H., et al. (2014). Constraining the recent mass balance of Pine Island and Thwaites glaciers, West Antarctica, with airborne observations of snow accumulation. The Cryosphere, 8, 1375–1392. https://doi. org/10.5194/tc-8-1375-2014

[RD88]	Michel, A., Flament, T., and Remy, F.: Study of the Penetration Bias of ENVISAT Altimeter Observations over Antarctica in Comparison to ICESat Observations, Remote Sens., 6, 9412–9434, https://doi.org/10.3390/rs6109412, 2014.
[RD89]	Motoyama, H. and Fujii, Y., 1999: Glaciological Data Collected by the 38th Japanese Antarctic Research, Expedition During 1997– 1998, Jare data reports, 28, National Institute of Polar Research, Tokyo.
[RD90]	Müller, K and 6 others (2010) An 860 km surface mass-balance profile on the East Antarctic plateau derived by GPR. Ann. Glaciol., 51(55), 1-8 (doi: 10.3189/172756410791392718)
[RD91]	Myers, M., Doran, P., & Myers, K., 2022: Valley-floor snowfall in Taylor Valley, Antarctica, from 1995 to 2017: Spring, summer and autumn. Antarctic Science, 34(4), 325-335. doi:10.1017/S0954102022000256
[RD92]	Newkirk, M.H., Brown, G.S., 1992: Issues related to waveform computations for radar altimeter applications. IEEE Trans. Antennas Propag. 40, 1478–1488. https://doi. org/10.1109/8.204738
[RD93]	Nitu, R., Y.-A. Roulet, M. Wolff, M. Earle, A. Reverdin, C. Smith, J. Kochendorfer, S. Morin, R. Rasmussen, K. Wong, J. Alastrué, L. Arnold, B. Baker, S. Buisán, J.L. Collado, M. Colli, B. Collins, A. Gaydos, H.-R. Hannula, J. Hoover, P. Joe, A. Kontu, T. Laine, L. Lanza, E. Lanzinger, GW Lee, Y. Lejeune, L. Leppänen, E. Mekis, J.-M. Panel, A. Poikonen, S. Ryu, F. Sabatini, J. Theriault, D. Yang, C. Genthon, F. van den Heuvel, N. Hirasawa, H. Konishi, K. Nishimura and A. Senese, 2018: WMO Solid Precipitation Intercomparison Experiment (SPICE) (2012 - 2015), Instruments and Observing Methods Report No. 131.
[RD94]	Noh, M. and Howat, I. M., 2015: Automated stereo-photogrammetric DEM generation at high latitudes: Surface Extraction with TIN-based Search-space Minimization (SETSM) validation and demonstration over glaciated regions, Gisci. Remote Sens., 52, 198–217.
[RD95]	Nolan, M., Larsen, C., and Sturm, M., 2015: Mapping snow depth from manned aircraft on landscape scales at centimeter resolution using structure-from-motion photogrammetry, The Cryosphere, 9, 1445–1463, https://doi.org/10.5194/tc-9-1445-2015.
[RD96]	Over, J.-S. R., Ritchie, A. C., Kranenburg, C. J., Brown, J. A., Buscombe, D. D., Noble, T., Sherwood, C. R., Warrick, J. A., and Wernette, P. A.: Processing coastal imagery with Agisoft Metashape Professional Edition, version 1.6—Structure from motion workflow documentation, Processing coastal imagery with Agisoft Metashape Professional Edition, version 1.6—Structure from motion workflow documentation, U.S. Geological Survey, Reston, VA, https://doi.org/10.3133/ofr20211039, 2021.
[RD97]	Overly, T. B., Hawley, R. L., Helm, V., Morris, E. M., and Chaudhary, R. N.: Greenland annual accumulation along the EGIG line, 1959–2004, from ASIRAS airborne radar and neutron-probe density measurements, The Cryosphere, 10, 1679–1694, https://doi.org/10.5194/tc-10-1679-2016, 2016.
[RD98]	Pälli, A., J.C. Kohler, E. Isaksson, J.C. Moore, J.F. Pinglot, V.A.Pohjola, and H. Samuelsson (2002), Spatial and temporal variability of snow accumulation using ground-penetrating radar and ice cores on a Svalbard glacier. Journal of Glaciology. 48(162): 417-424.
[RD99]	Partington, K., Ridley, J., Rapley, C., Zwally, H., 1989: Observations of the surface properties of the ice sheets by satellite radar altimetry. J. Glaciol. 35, 267–275. https://doi.org/10.3189/S0022143000004603.
[RD100]	Picard, G., L. Brucker, M. Fily, H. Gallée, and G. Krinner, 2009: Modeling time series of microwave brightness temperature in Antarctica. J. Glaciol., 55, 537–551.
[RD101]	Picard, G., L. Arnaud, J.-M. Panel, and S. Morin, 2016: Design of a scanning laser meter for monitoring the spatio-temporal evolution of snow depth and its application in the Alps and in Antarctica. The Cryosphere, 10, 1495–1511.
[RD102]	Picard, G., Sandells, M., Lowe, H., 2018: Smrt: an active–passive microwave radiative transfer model for snow with multiple microstructure and scattering formulations (v1.0). Geosci. Model Dev. 11, 2763–2788. https://doi.org/10.5194/gmd-11-2763- 2018.

[RD103]	Picard, G., L. Arnaud, R. Caneill, E. Lefebvre, and M. Lamare, 2019: Observation of the process of snow accumulation on the Antarctic Plateau by time lapse laser scanning. The Cryosphere, 13, 1983–1999.
[RD104]	Prokop, A., 2008: Assessing the applicability of terrestrial laser scanningfor spatial snow depth measurements, Cold Reg. Sci. Technol., 54, 155–163.
[RD105]	Qin, D., Xiao, C., Allison, I., Bian L., 2004: Stephenson, R., Ren, J., and Yan, M.: Snow surface height variations on the Antarctic ice sheet in Princess Elizabeth Land, Antarctica: 1 year data from an automatic weather station, Ann. Glaciol., 39, 181–187.
[RD106]	Reijmer, C. H., and M. R. V. D. Broeke, 2003: Temporal and spatial variability of the surface mass balance in Dronning Maud Land, Antarctica, as derived from automatic weather stations, J. Glaciol., 49(167), 512–520.
[RD107]	Rémy, F., Parouty, S., 2009: Antarctic ice sheet and radar altimetry: a review. Remote Sens. 1, 1212–1239. https://doi.org/10.3390/rs1041212.
[RD108]	Remy, F., Flament, T., Blarel, F., Benveniste, J., 2012: Radar altimetry measurements over antarctic ice sheet: a focus on antenna polarization and change in backscatter problems. Adv. Space Res. 50, 998–1006. https://doi.org/10.1016/j. asr.2012.04.003.
[RD109]	Retzlafl, R., Lord, N. and Bentley, C.R.. 1993. Airborne radar studies: lee Streams A, Band C, West Antarctica. J.Glacial., 39(133), 495–506 10.1017/S0022143000016397
[RD110]	Revuelto, J., Alonso-Gonzalez, E., Vidaller-Gayan, I., Lacroix, E., Izagirre, E., Rodríguez-López, G., and López-Moreno, J. I., 2021: Intercomparison of UAV platforms for mapping snow depth distribution in complex alpine terrain, Cold Regions Science and Technology, 190, 103344, https://doi.org/10.1016/j.coldregions.2021.103344.
[RD111]	Rodríguez-Morales, F and 9 others (2014) Advanced multifrequency radar instrumentation for polar research. IEEE Trans Geosci. Remote Sens., 52(5), 2824–2842
[RD112]	Ryan, W. A., Doesken, N. J., and Fassnacht, S. R., 2008: Evaluation of Ultrasonic Snow Depth Sensors for U.S. Snow Measurements, J. Atmos. Ocean. Tech., 25, 667–684.
[RD113]	Rosmorduc, V., Benveniste, J., Lauret, O., Maheu, C., Milagro, M., Picot, N., 2011: Radar Altimetry Tutorial. http://www.altimetry.info/radar-altimetry-tutorial/ (Last access: October 2020).
[RD114]	Seefeldt, M. W., Low, T. M., Landolt, S. D., and Nylen, T. H., 2021: Remote and autonomous measurements of precipitation for the northwestern Ross Ice Shelf, Antarctica, Earth Syst. Sci. Data, 13, 5803–5817, https://doi.org/10.5194/essd-13-5803-2021.
[RD115]	Shepherd, A., and Coauthors, 2012: A Reconciled Estimate of Ice-Sheet Mass Balance. Science, 338, 1183–1189.
[RD116]	Slater, T., Shepherd, A., McMillan, M., Muir, A., Gilbert, L., Hogg, A. E., Konrad, H., and Parrinello, T., 2018: A new digital elevation model of Antarctica derived from CryoSat-2 altimetry, The Cryosphere, 12, 1551–1562.
[RD117]	Smith, M. W., 2014: Roughness in the Earth Sciences, Earth Science Reviews, 136, 202-225.
[RD118]	Smith, M.W., Vericat, D., 2015: From experimental plots to experimental landscapes: topography, erosion and deposition in sub-humid badlands from structure-from motion photogrammetry. Earth Surf. Process. Landf. 40, 1656–1671. https://doi. org/10.1002/esp.3747.
[RD119]	Sørensen, L.S., S.B. Simonsen, R. Forsberg, K. Khvorostovsky, R. Meister, M.E. Engdahl, 25 years of elevation changes of the Greenland Ice Sheet from ERS, Envisat, and CryoSat-2 radar altimetry, Earth and Planetary Science Letters, 495, 234-241, 2018.
[RD120]	Smith, M.W., Quincey, D.J., Dixon, T., Bingham, R.G., Carrivick, J.L., Irvine-Fynn, T.D.L., Rippin, D.M., 2016: Aerodynamic roughness of glacial ice surfaces derived from high-resolution topographic data. J. Geophys. Res. 121, 748–766. https://doi.org/ 10.1002/2015JF003759.Received.

[RD121]	Trouvilliez, A., F. Naaim-Bouvet, H. Bellot, C. Genthon, and H. Gallée, 2015: Evaluation of the FlowCapt Acoustic Sensor for the Aeolian Transport of Snow. J. Atmospheric Ocean. Technol., 32, 1630–1641.
[RD122]	Van den Broeke, M. R., C. H. Reijmer, and R. S. W. Van De Wal, 2004: A study of the surface mass balance in Dronning Maud Land, Antarctica, using automatic weather stations. J. Glaciol., 50, 565–582.
[RD123]	Vander Jagt, B., Lucieer, A., Wallace, L., Turner, D., and Durand, M., 2015: Snow Depth Retrieval with UAS Using Photogrammetric Techniques, Geosciences, 5, 264–285, https://doi.org/10.3390/geosciences5030264.
[RD124]	Van Der Veen, C.J., Krabill, W.B., Csatho, B.M., Bolzan, J.F., 1998: Surface roughness on the Greenland ice sheet from airborne laser altimetry. Geophys. Res. Lett. 25, 3887–3890. https://doi.org/10.1029/1998GL900041.
[RD125]	Van Der Veen, C.J., Ahn, Y., Csatho, B.M., Mosley-Thompson, E., Krabill, W.B., 2009: Surface roughness over the northern half of the Greenland Ice sheet from airborne laser altimetry. J. Geophys. Res. 114, 1–8. https://doi.org/10.1029/2008JF001067.
[RD126]	van Lipzig, N. P. M., J. Turner, S. R. Colwell, and M. R. van den Broeke, 2004, The near-surface wind field over the Antarctic continent, Int. J. Climatol., 24(15), 1973–1982.
[RD127]	Verfaillie, D., Fily, M., Le Meur, E., Magand, O., Jourdain, B., Arnaud, L., and Favier, V.: Snow accumulation variability derived from radar and firn core data along a 600 km transect in Adelie Land, East Antarctic plateau, The Cryosphere, 6, 1345–1358, https://doi.org/10.5194/tc-6-1345-2012, 2012.
[RD128]	Vignon, E., and Coauthors, 2017: Stable boundary-layer regimes at Dome C, Antarctica: observation and analysis. Q. J. R. Meteorol. Soc., 143, 1241–1253.
[RD129]	Walford, M.E.R. (1964), Radio-echo sounding through an ice shelf. Nature. 204: 317-319.
[RD130]	Welch, B.C., W.T. Pfeffer, J.T. Harper and N.F. Humphrey (1998), Mapping subglacial surfaces of temperate valley glaciers by two-pass migration of radio-echo sounding survey data. Journal of Glaciology. 44: 164-170.
[RD131]	Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., and Roth, A., 2018: Accuracy Assessment of the Global TanDEM-X Digital Elevation Model with GPS Data, ISPRS J. Photogramm., 139, 171–182.
[RD132]	Wingham, D.J., C.R. Francis, S. Baker, C. Bouzinac, D. Brockley, R. Cullen, P. de Chateau-Thierry, S.W. Laxon, U. Mallow, C. Mavrocordatos, L. Phalippou, G. Ratier, L. Rey, F. Rostan, P. Viau, D.W. Wallis, 2006: CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields, Advances in Space Research, 37 (4), 841-871.
[RD133]	Winebrenner, D.P., B.E. Smith, G.A. Catania, H.B. Conway, and C.F. Raymond (2003), Radio-frequency attenuation between Siple Dome, West Antarctica, from wide-angle and profiling radar observations. Annals of Glaciology. 37: 226-232.
[RD134]	Whitehouse, P. L., N. Gomez, M. A. King, and D. A. Wiens, 2019: Solid Earth change and the evolution of the Antarctic Ice Sheet. Nat. Commun., 10, 503.
[RD135]	Yi, D., Bentley, C., 1994: Analysis of satellite radar-altimeter return wave forms over the east antarctic ice sheet. Ann. Glaciol. 20, 137–142. https://doi.org/10.3189/ 1994AoG20-1-137-142.
[RD136]	Young, D., Lindzey, L., Blankenship, D., Greenbaum, J., Garcia De Gorordo, A., Kempf, S., et al., 2015: Land-ice elevation changes from photon-counting swath altimetry: First applications over the Antarctic ice sheet. Journal of Glaciology, 61(225), 17-28. doi:10.3189/2015JoG14J048
[RD137]	Zwally, H. J., Schutz, R., Hancock, D., and DiMarzio, J., 2014: GLAS/ICESat L2 Antarctic and Greenland Ice Sheet Altimetry Data (HDF5), Version 34. Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center.
[RD138]	Zwally, H. J., J. Li, J. W. Robbins, J. L. Saba, D. Yi, and A. C. Brenner, 2015: Mass gains of the Antarctic ice sheet exceed losses. J. Glaciol., 61, 1019–1036.

[RD139]	Zwally, H. J., J. Li, J. W. Robbins, J. L. Saba, D. Yi, and A. C. Brenner, 2015: Mass gains of the Antarctic ice sheet exceed losses. J. Glaciol., 61, 1019–1036.

Table of contents

1 Introduction

1.1 Purpose and scope

This document is the WP 3.1 FRM Protocols and Procedures for S3 STM Land Ice Products for the "Sentinel-3 Topography mission Assessment through Reference Techniques (St3TART)" project, [RD1].

This deliverable presents a review of the state-of-the-art existing in-situ solutions, procedures, methods, fiducial reference measurement (FRM) data and associated uncertainties that have been already used to assess the performance of satellite altimetry over land ice. This is done to ensure that these observations meet the criteria of FRM standards and can be used in an operational way for the validation of the Sentinel-3 Land topography mission.

1.2 Overview of this document

In addition to this Introduction chapter, this "FRM Protocols and Procedures" document includes the following chapters:

- ▲ Chapter 2: The altimetry measurement for land ice
- ▲ Chapter 3: The Cal/Val framework for land ice
- ▲ Chapter 4: Requirements for Cal/Val altimetry measurements for land ice
- ▲ Chapter 5: Existing means for Cal/Val activities
    - o State of the art of existing sensors with ranking and discussion on their FRM compliancy
    - o Existing networks, surveys and campaigns
- ▲ Chapter 6: Use of other Satellite missions
- ▲ Chapter 7: FRM Protocols and Procedures

1.3 Definition of FRM and metrological approach

The Quality Assurance framework for Earth Observation (QA4EO), established and endorsed by Committee on Earth Observation Satellites (CEOS), defines the following principle regarding Earth Observation data quality:

> *'It is critical that data and derived products are easily accessible in an open manner and have an associated indicator of quality traceable to reference standards (preferably SI) so users can assess suitability for their applications i.e., 'fitness for purpose'.'*

QA4EO defines high level processes to achieve this, such as well-documented procedures, participation in comparisons and uncertainty assessments that apply to all EO data records. Traceability requires that this quality indicator be based on "a documented and quantifiable assessment of evidence demonstrating the level of traceability to internationally agreed (where possible SI) reference standards." The QA4EO principle stops short of requiring SI-traceability in all circumstances, recognizing that the full rigor of linkage to SI may not be viable for all applications and measurements, however, the accompanying guidelines are based on metrological concepts adapted from guidelines of the international metrology community and a metrological approach is strongly implied.

Metrology, the science of measurement, is the discipline responsible for maintaining the International System of Units (SI) and the associated system of measurement. It is core to the SI that measurements are stable over centuries and that measurement standards are equivalent worldwide. These properties are achieved through the key principles of metrological traceability: uncertainty analysis and comparison. These same principles are important to Earth observation.

The term "Fiducial Reference Measurement" (FRM) is used for non-satellite observations that follow QA4EO guidelines.

> *Fiducial reference measurements (FRMs) are a suite of independent, fully characterized, and traceable sub-orbital measurements that follow the guidelines outlined by the GEO/CEOS Quality Assurance framework for Earth Observation (QA4EO) and have value for space-based observations.*

Thus, FRMs are the quality-assured non-satellite observations that can be used to calibrate and validate satellite-based sensor measurements. As ESA states 'these FRM provide the maximum return on investment for a satellite mission by delivering to users, the required confidence in data products, in the form of independent validation results and satellite measurement uncertainty estimation, over the entire end-to-end duration of a satellite mission.'

Other satellite observations cannot be considered FRMs under this definition, but satellite-based comparisons are important in the calibration and validation of satellite missions and to ensure the continuity of the time series between the different missions, in particular for the analysis of climatic evolutions and to feed the reanalysis models. These multi-mission comparisons also make it possible to estimate the coherence and stability of space observations despite the change of space support or sensor types. Thus, a satellite-based comparison should also be considered metrologically, following many similar approaches. Similarly, comparisons with model outputs (e.g. reanalyses) are also used in Cal/Val processes, but are not FRMs.

2 The altimetry measurement for land ice

2.1 Altimetry applications

Satellite altimetry over land ice is mainly used for two complementary purposes:

▲ To create digital surface models (DSM) by interpolation between ground-tracks or by using altimetry data to calibrate the positioning of other satellite DSM techniques like photogrammetry and SAR interferometry;

▲ To repeatedly measure surface elevation (or a near-surface reference horizon) so precisely that ice thickness changes can be detected and mapped in space and time.

These two main applications put different demands on the satellite data. For the first purpose, a widespread coverage of data is needed over a relatively short amount of time, so that the snow/ice surface has not changed in the meantime. Sentinel-3 has to our knowledge not been used for DSM applications over land ice yet, but the cryosphere-designated altimeters CryoSat-2 and ICESat/ICESat-2 have been widely used to make DSMs directly over the ice sheets through interpolation between ground-tracks (Helm et al. 2014; Slater et al. 2019) and as horizontal and vertical reference for satellite photogrammetry DSMs such as the widely used models GIMP for Greenland (Howat et al. 2014), REMA for Antarctica (Howat et al. 2019) and ArcticDEM for all Arctic regions (Porter et al. 2018). Considering the ongoing improvements that are being made to the Sentinel-3 operation, processing, and thematic data products, it is likely that these data will soon also be used for DSM applications at high latitudes where track spacing is sufficiently dense.

For the second purpose, it is important to monitor the same locations repeatedly over long periods of time. Surface elevation over land ice evolves in response to variations in surface mass balance (SMB, the net of surface accumulation and ablation, e.g., Agosta et al. 2019), firn densification (e.g, Ligtenberg et al. 2011), ice dynamics (e.g., Edwards et al. 2019), and to a smaller extent basal melting and Glacial Isostatic Adjustment (GIA, e.g., Whitehouse et al. 2019). Future changes in snow accumulation, ablation (mainly melting) and ice dynamics due to climate change will be directly transferred into surface elevation changes. Satellite altimetry is the most precise tool to measure land-ice elevation changes (e.g., Bamber et al. 2009; Shepherd et al. 2012; IMBIE team 2018) and through that gain knowledge of these processes which is crucial for society to assess impacts of climate change, in particular sea level rise from land-ice loss.

2.2 Altimetry principles

A radar altimeter measures the return signal from emitted radar pulses that get reflected by Earth's surface. Precise satellite-to-surface ranges are derived by identifying the surface reflection point in the return radar waveforms. When combined with precise positioning of the satellite vehicle, this allows estimation of along-track surface elevation through time. The altimetry system onboard Sentinel-3 consists of a dual-frequency (C-band/Ku-band) SAR Radar Altimeter (SRAL), a Microwave Radiometer (MWR) and a Precise Orbit Determination (POD) system which further includes a GNSS instrument for positioning wrt. higher orbiting GNSS satellites, a Doppler Orbit determination and Radio-positioning instrument (DORIS) that for positioning wrt. DORIS ground stations, and a passive Laser Retro-Reflector for precise laser tracking of the satellite from the ground. The C-band SRAL and MWR instruments are primarily used for dual-frequency corrections of ionospheric and tropospheric signal delays over the ocean and are not applied over land surfaces where these corrections are instead retrieved form a Global Ionosphere Model (GIM, Iijima et al., 1999) and from ECMWF meteorological data for dry and wet troposphere corrections [RD8].

A major benefit with Sentinel-3 SRAL compared to conventional radar altimeters is its delay-Doppler SAR capability that allows improved along-track resolution by combining a series of consecutive radar pulses (4 bursts of 64 pulses, 256 in total) into a multi-looked waveform with an along-track footprint of only ~330 m (Fig. 2.1). The across-track footprint is still around 15 km, as for the typical Low Resolution Mode (LRM) altimeters, but careful analysis of the multi-looked waveform in combination with a-priori knowledge of surface topography makes it possible to locate the initial reflection point within the cross-track beam at an accuracy of a few hundred metres or better. The potential accuracy of derived surface elevations follows those of the satellite positioning and the surface ranging, both at the centimetre scale [RD7], but this is difficult to achieve over areas with surface topography due to its complex impact on the return waveform and the consequent problem of locating where the backscattered energy comes from within the SAR footprint.
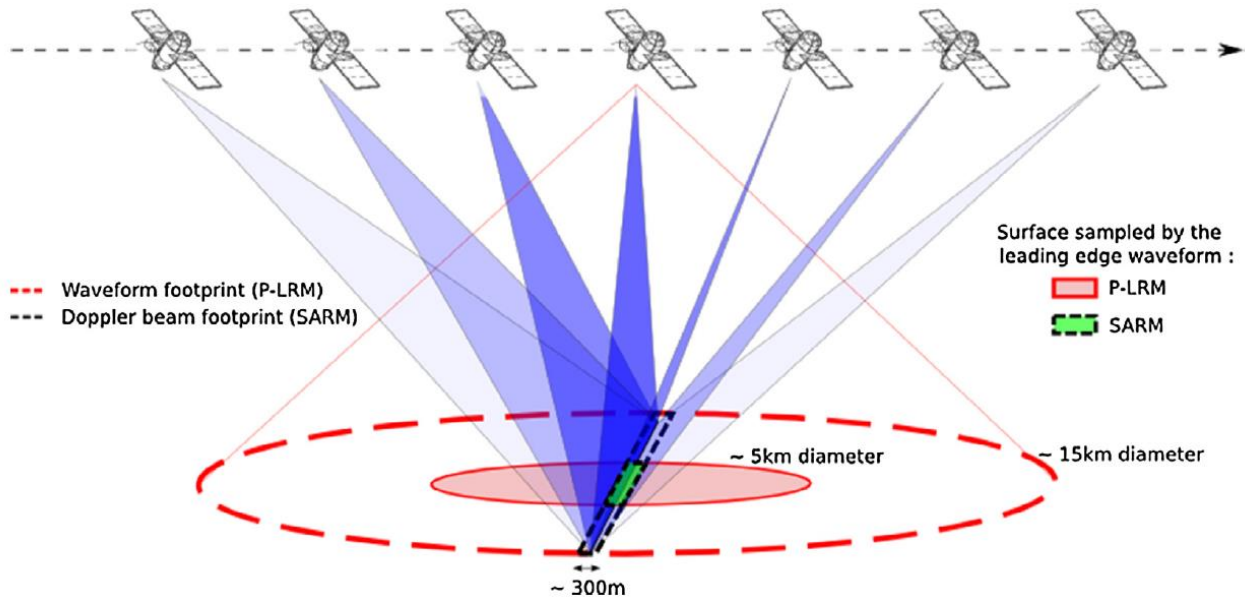
Figure 2.1 Conceptual sketch of a SAR-mode altimeter (Aublanc et al., 2018) where a series of along-track radar looks are combined by delay-Doppler processing into a narrow footprint with a single stacked waveform. In contrast, a Low Resolution Mode (LRM) altimeter uses one radar pulse resulting in a circular ~15 km footprint. The actual area that contributes to the leading edge of the waveform, where surface elevation is retrieved, is considerably smaller. Sentinel-3 has only a SAR altimeter, but the data can also be processed as pseudo-LRM observations.

Since the relative area contribution to the return waveform in the outer parts of the footprint is greatly reduced in SAR mode, the waveform appears sharper and peakier than in LRM mode (Fig. 2.2). This is true for both the surface and sub-surface (within the snowpack) contribution to the waveform, and the surface point can be estimated from a relatively narrow zone of the leading edge. This process is known as retracking and relies on empirical algorithms or physical models that are fit to the waveforms. The standard Sentinel-3 processing over land-ice employs the OCOG/ICE-1 retracker (Wingham et al., 1986; Bamber, 1994) and the UCL ice-sheet retracker (Wingham and Wallis, 2010), where the first one uses a pre-defined threshold applied to a calculated amplitude, and the second one fits a gaussian-shaped echo model to the waveform which is optimized for interior ice-sheet surfaces with low slopes.
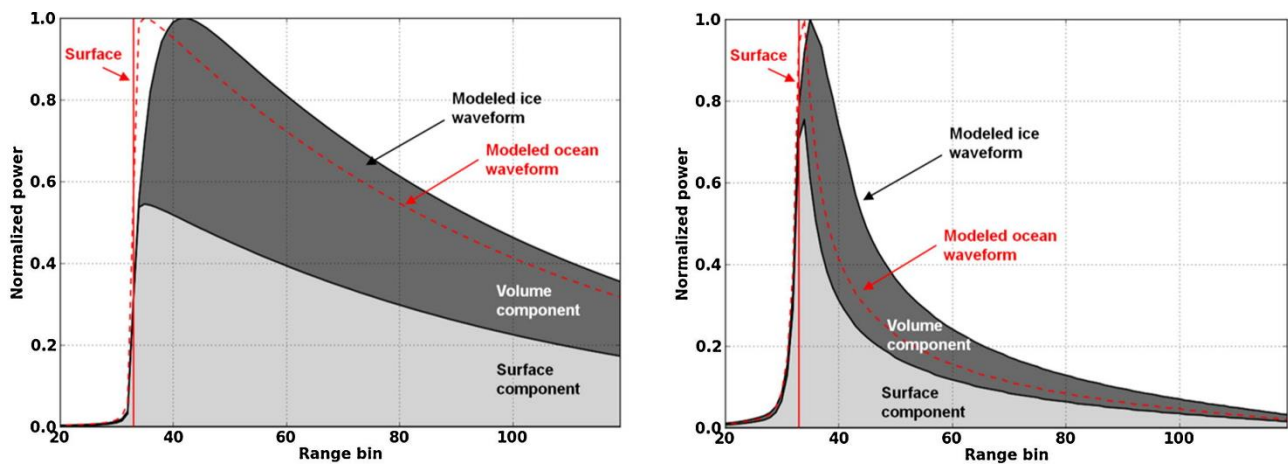
Figure 2.2 Idealized models of the return waveforms for LRM altimetry (left) and SAR altimetry (right) over the ocean and over flat land ice (Aublanc et al., 2018), illustrating the peakier nature of a SAR waveform due to its smaller footprint, as well as the impact of volume scattering in snow (dark grey) compared to the pure surface echo (light grey). The satellite range to the surface is derived from the leading edge of the waveform (red line) using a designated re-tracking algorithm for the relevant surface type.

Ice sheets have complex terrains with significant surface undulations and slopes, varying from gentle slopes and variable roughness in the inland to higher slopes and rougher terrain along the coast, sometimes including a flatter floating part known as ice shelves (Fig. 2.3). Over such surfaces, the returned waveforms can have a complex set of peaks. Most commonly, the first reflection is used to identify the 'point of closest approach' (POCA), but there can also be peaks further out in the waveform that can be resolved if the topography is known or can be modelled without ambiguities. Here, we will focus on POCA as this is the basic principle for the observational relocation in the Sentinel-3 thematic product for land-ice. The geolocation of POCA will depend on the surface topography and can deviate by several kilometres from the satellite nadir ground track (Fig. 2.4). Since the satellite orbits can vary with a few hundred meters between different overpasses, the repeated POCA tracks can deviate even more due to the additional impact of variable topography. This makes it challenging to validate POCA data and to determine surface elevation changes over time.
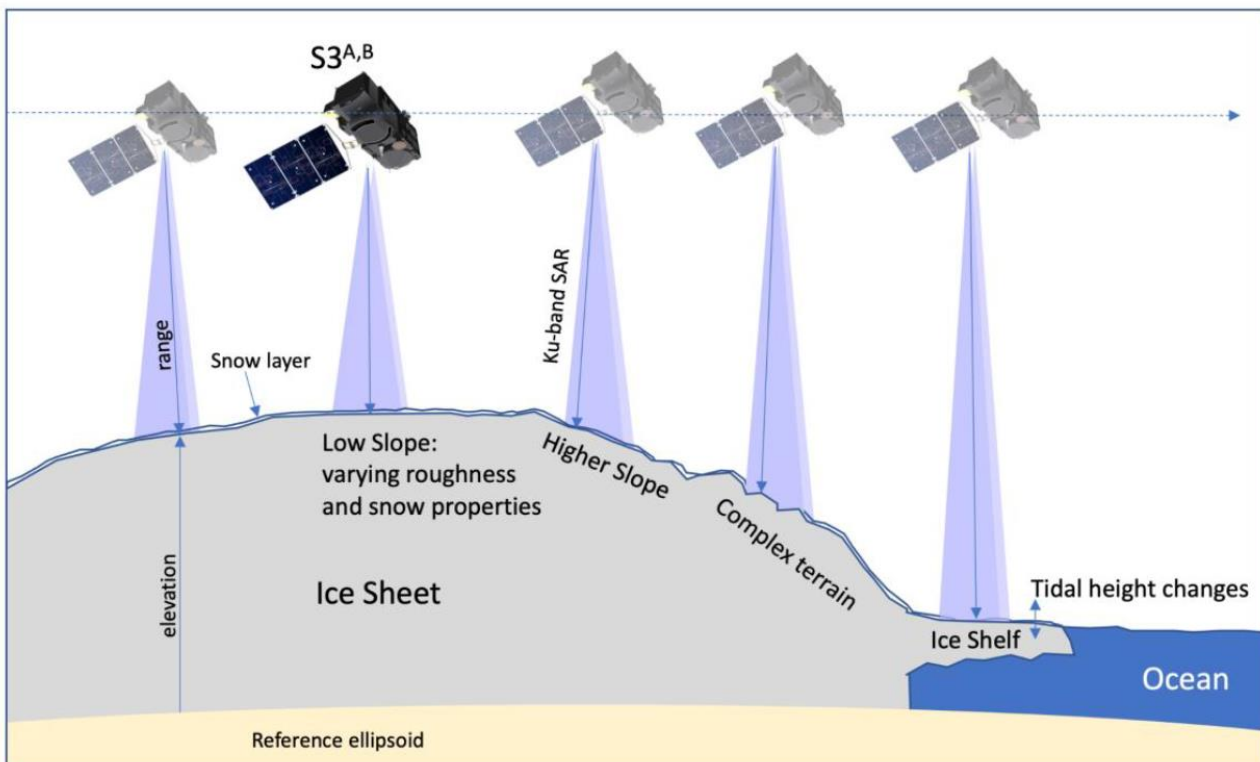
Figure 2.3 Sentinel-3 altimetry over the various surface types of an ice sheet (credit: Mullard Space Science Laboratory, MSSL)

Ideally, a high-resolution DSM would be used to geolocate POCA observations precisely within each footprint, but in practice this is very computing intensive and not straightforward to do, as illustrated by Figure 2.4. ESA's Instrument Processing Facility (IPF) for the land-ice thematic product has implemented a linear slope model for POCA relocation in Antarctica and Greenland based on two similar DSMs made from CryoSat-2 altimetry (Helm et al., 2014). This works well in areas where the footprint-scale topography can be approximated by a linear surface, but that is often not the case in coastal parts of Antarctica and Greenland where sharp transitions can occur around nunataks, outlet glaciers, and the edges of ice shelves. Polar ice caps and continental mountain glaciers are even more challenging, and as of March 2023 there is no POCA relocation applied to the land-ice thematic products in these regions. For a further discussion of these issues, we refer to the auxiliary technical report to this document [RD10] which investigates the added value of high-resolution DEMs for processing and interpretation of Sentinel-3 altimetry data.

Over an ice shelf, the surface is flatter and POCA deviations are smaller. However, the ice surface is floating on water. This means that only around 10 % of the ice is above water, and each 1 m change in the ice thickness will only be seen as a 0.1 m change in the measured surface height (e.g. Howat et al., 2019). Because of this, observations require an even higher accuracy for glaciological applications here, and additional corrections need to be considered for ocean tides and the inverse barometer effect (IBE) on the ocean due to surface pressure variations. Measurement equations, corrections and uncertainty diagrams for Sentinel-3 thematic data products over grounded ice and ice shelves are presented in Section 7, along with how these data can be compared with FRMs in space and time.
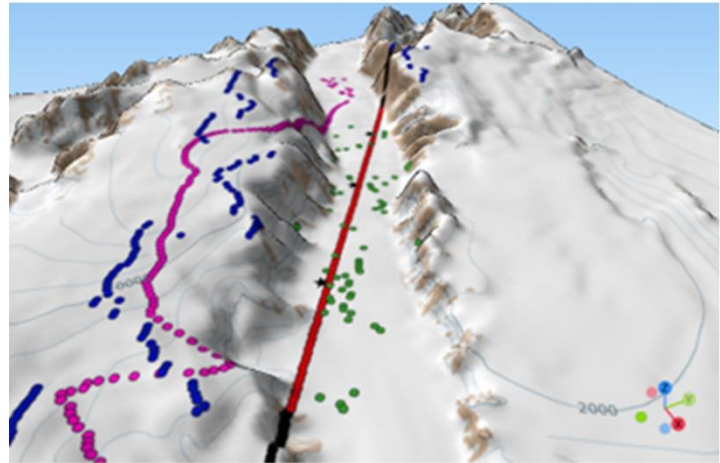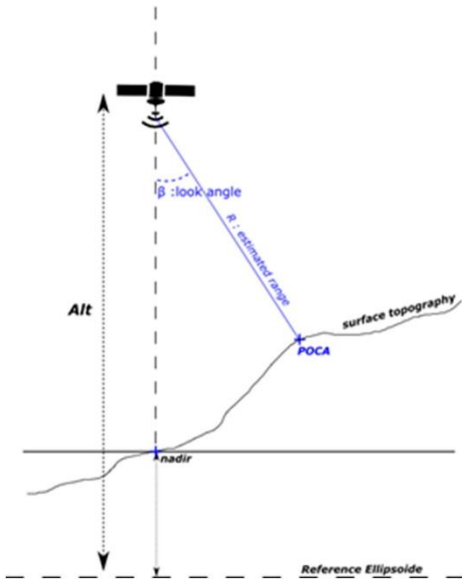
Figure 2.4 Cross-track illustration of the satellite point of closest approach (POCA) for altimetry observations (left panel) and how relocated observations on the ground (magneta for ESA-IPF, blue for Roemer et al., and green for CLS) can vary from satellite nadir (red) over a valley glacier (right panel). This is further discussed in the auxiliary technical report [RD10] where this figure is extracted from (credit J. Aublanc).

A further challenge over rough terrain is the ability of the satellite altimeter to rapidly adjust its sampling window to the actual timing of the surface returns. Sentinel-3 has the ability to alternate between a closed-loop tracking command (CLTC) where the window is adjusted on-the-fly according to the onboard surface tracker, and an open-loop tracking command (OLTC) where the window is adjusted according to along-track topography data extracted from a DSM and uploaded to the satellite (Le Gac et al., 2019). Since October 2020, OLTC has been implemented for all mid-latitude continental areas [RD8] and is being tested for selected cycles and regions in polar areas, including the ice sheets.

3 Calibration-Validation framework for land ice

3.1 Special considerations for land ice

3.1.1 Land ice as a measurand

To assess land-ice thickness changes and mass balance, we need to assume that the bedrock surface underneath the ice is stable or that its uplift (or subsidence) can be corrected. We also need a stable reference point for the upper surface which can be the true surface of the snow (or ice in case of no snow) or a sub-surface horizon represented by a physical transition in material structure or density. In terms of a calibration-validation framework, it can be helpful to consider the surface as an underlying structure, with smaller spatial and temporal structure on top of this. A simplified cross-section of such a surface is visualized in Figure 3.1 with layers consisting of:

- *Snow* that has recently accumulated and in dry/cold conditions being eroded and deposited in new areas by wind, forming sastrugi and dunes. Snow crystals are further reshaped by dry or wet metamorphism processes.
- *Firn* or consolidated snow, that is a compressed, transformed and coarser stage of snow that can be considered more stable than the overlying fresh snow layer. In regions with significant summer snow melt, the firn layer has distinctly different properties due to the impact of liquid water and refreezing, whereas in all-year cold regions like the interior of the ice sheets, the transition is diffuse but sometimes detectable by snow pit/core analysis or radargrams.
- *Ice* that is compressed firn over many decades on ice sheets or over only a few years on glaciers and ice caps where high temperatures and internal refreezing of meltwater speeds up the transformation from firn to ice. Ice can also be exposed at the surface in areas where snow and firn layers have been eroded by wind or melted away, for example in the blue ice areas of Antarctica and for glacier fronts in the Arctic where surface mass balance is negative and the ice originates from the inland.

All these layers move with the ice flow, either uniformly by basal sliding (typically tens of metres per year) or more slowly (a few metres per year) from internal deformation in cases where the bed is frozen to the ground.
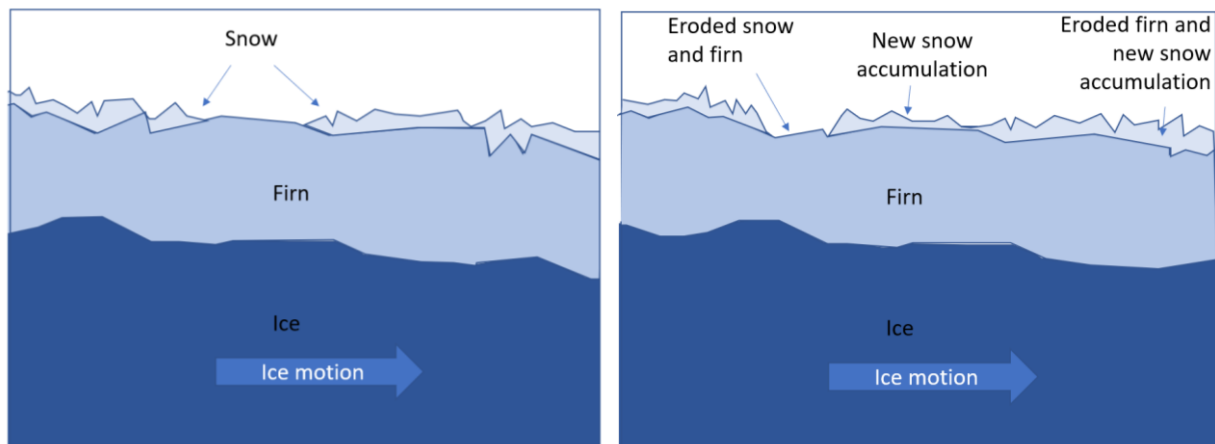
Figure 3.1 Cross section of the land-ice surface showing motion in the relative changes between the left and right figure.

Over time, new snow accumulates, and snow and firn erode and deposit in new places. In milder climates, snow melts down and sometimes expose firn or ice at the surface in summer. All this means that the surface is changing dynamically, so there is no simple reference surface for ground-based validation measurements. Surface elevation changes are also difficult to assess due to lack of stable reference elevations or ground control points to monitor the change against.

Regions with strong snow drift have a particularly complex structure with depositional features at various scales from smaller sastrugi to massive dunes (Fig. 3.2). These formations evolve and migrate with time and align in various ways with the dominant wind direction. The drifting snow provides a very dynamic surface, with spatial scales of decametres that varies in hours, and sometimes minutes. Any in situ measurement needs to average over sufficient spatial and temporal scale to minimise the effect of the drifting snow structure. It also needs to be updated regularly enough to account for the other slower-process changes such as firn compaction and ice motion.
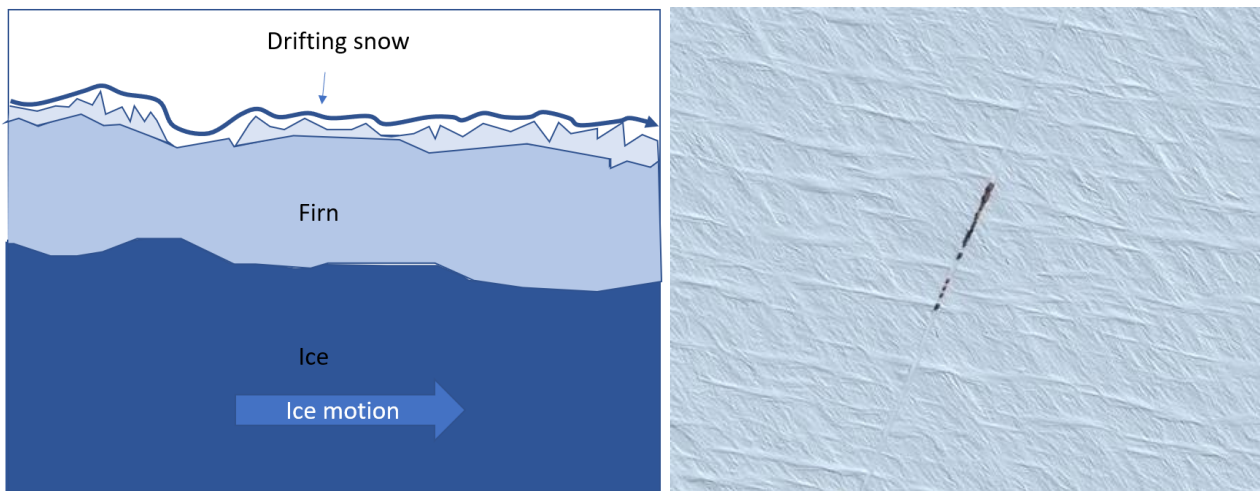
Figure 3.2 Left: cross section of surface shown with additional drifting snow layer. Right: SPOT-6 satellite image of an Antarctic snow area showing two different scales of surface roughness: sastrugi (~10s cm high, ~1-5 m long) and dunes (50-100 cm high, 200 m long). The line of black squares are snow vehicles and sledges along the logistics traverse route supplying the Dome Concordia station on the Antarctic inland plateau.

A more detailed discussion of the land-ice surface as a measurand is provided in Annex A (Section 8).

3.1.2   Signal penetration in snow

Depending on the further use of the altimetry data over land ice, the measurand might have to be corrected for signal interaction within the near-surface medium. The radar altimeter signal can be biased from the surface of cold/dry snow, because the wave partially interferes with the irregular surface and with the snowpack (Partington et al., 1989; Legrésy and Rémy, 1997; Rémy and Parouty, 2009). This is comparable to the sea-state effect, a well-known problem over the ocean. Larue et al. (2021) accurately describes how the radar waveform is resulting from the echo coming from the upper air-snow interface ("surface echo"), and the diffuse echo caused by the reflections at internal layer interfaces within the snow pack ("internal interface echoes") and by scattering by snow grains ("volume echo") (Brown, 1977, Partington et al., 1989). A model of the surface and volume contribution to a return waveform is shown in Figure 2.2.

The surface echo is mainly controlled by the topography that influences the illuminated radar footprint (Brenner et al., 1983), by the surface roughness (of centimetre to decimetre scale) that causes backscatter differences between rough (weak backscatter) and smooth (stronger backscatter) surfaces (Van Der Veen et al., 1998; Van Der Veen et al., 2009; Rosmorduc et al., 2011; Kurtz et al., 2014; Smith and Vericat, 2015; Smith et al., 2016), and by the surface density (strength of the reflection). The volume echo depends on the penetration of the radar wave, which can be a few metres into the snowpack at high frequencies like Ku-band (13.6 Ghz for S3 SRAL) but is deeper at lower frequencies like C-band or L-band (Legrésy and Rémy, 1997). The echo depends on scattering mechanisms in the snowpack (Remy et al., 2012), which are themselves a function of grain size, wave frequency, and internal stratification of the snowpack based on snow density contrasts between successive layers. The echo also results from signal absorption losses that depend on snow temperature and wave frequency primarily (Adodo et al., 2018) inducing visible seasonal variations. The volume echo is therefore closely related to the properties of the snow. As these are extremely variable at the scale of glaciers and ice sheets, the precise quantification of the variations of the volume echo is complex over the continent but can be characterized thanks to the combination of altimetry observations at multiple frequencies Ka/Ku bands (Guerreiro et al., 2016, 2017) or by cross-comparison with laser altimetry which has minimal signal penetration (Smith et al. 2018).

Several waveform models have been developed to investigate the radar signal over the ice sheets (Newkirk and Brown, 1992; Femenias et al., 1993; Adams and Brown, 1998). These models generally assume a homogeneous snowpack. However, Larue et al., (2021) proposed to extend the Snow Model Radiative Transfer model (SMRT, Picard et al., 2018) to conduct waveform simulations on the Antarctic ice sheet with more realistic snowpacks. They investigated how the penetration depth of the wave depends on snow properties and estimated the induced elevation bias. Their main finding is that this bias is relatively independent of the frequency, as opposed to past affirmation in the literature, mainly because despite a lower penetration at higher frequencies, the volume echo is (proportionally) larger, returning more signal from the volume of the snowpack rather than from the surface.

Another factor to consider is the impact of surface melting, percolation in the snow/firn pack and potential refreezing into ice lenses and layers within the snowpack. The onset of surface melt (or rainfall) can cause a rapid transition from a dry and relatively homogeneous snowpack to one with heterogenous refrozen ice layers near the surface, with a potential shift from volume-dominant to a surface-dominant radar backscattering. This has been observed to happen as high as the summit of Greenland, where an extreme melt event in July 2012 was followed by an apparent elevation increase of half a metre in CryoSat-2 Ku-band altimetry, attributed to a sudden rise in the reflective surface (Nilsson et al. 2015). This is a serious concern for altimetry-derived time series of elevation change in areas where surface melting (or rain) happens sporadically, whereas in milder climates where melting occurs every summer, it is mainly causing a seasonal disturbance in the signal, which is not affecting the long-term rate of change (Morris et al. 2021). As snow builds up during the cold winter season, the volume component of the return waveform will again increase, and this has been used to measure snow accumulation (depth) over the last summer surface (top of firn) using airborne Ku-band radar (Hawley et al. 2013). However, they noted that this would be difficult to achieve with spaceborne radar altimeters due to the coarser range resolution (~47 cm for Sentinel-3) and the larger footprint area contributing to the echo.
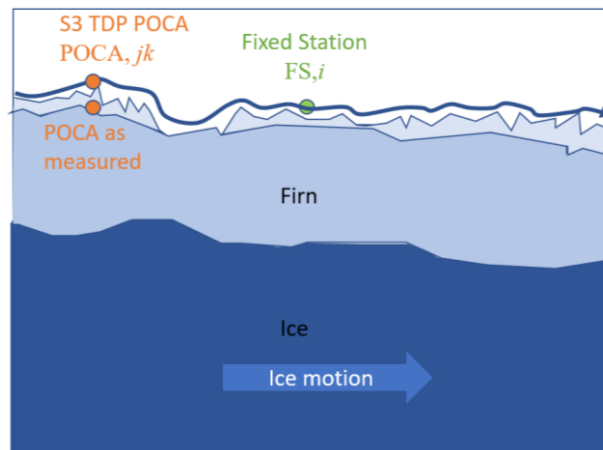
Figure 3.3 The Sentinel-3 thematic data product (TDP) for land ice is referenced to the surface at the point of closest approach (POCA), but depending on snow/firn properties and applied retracker, POCA as measured can be below the actual surface due to radar penetration.

A complete FRM system should consider both the all-year cold and dry volume scattering regime in the interior of the ice sheets and the seasonal melt/freeze regime of glaciers and coastal parts of the ice sheet, including the transitional zones where extreme events of melt or rain occasionally occur, intensified by climate change. A fixed FRM station should also collect auxiliary data on meteorology and snowpack properties, for example with a densely spaced thermistor string at the air-snow interface where periods of melt and refreezing can be monitored. This is discussed in Section 5.3.5.

3.1.3   Retracking and georeferencing of POCA

As described in Section 2, the leading-edge of the return waveform stems from the point of closest approach (POCA, Fig. 2.4) which is the reference for retracking and geolocation the thematic data product over the ice sheets, although POCA relocation is not yet implemented for glacier regions outside of Antarctica and Greenland (as of March 2023). POCA locations are estimated from a linear slope model derived from an ice sheet DSM (Helm et al., 2014), but despite a stable reference topography model, repeat-pass POCA tracks vary considerably (hundreds of metres to kilometres) due to the topographic enhancement of small deviations from the reference orbits. A more serious concern is that the actual POCA location, where the leading-edge of the echo comes from in reality, might be equally different due to errors in the topography model or the simplification of assuming linear slopes in areas where there is substantial surface roughness or other topographic irregularities.

Variations and issues in POCA retracking and geolocation are most severe for the coastal parts of the ice sheet where there is a variety of surface topography types, from relatively flat and homogeneous ice shelves to rough ones with ridges and rifts, and from smooth but steep ice rises and promontories to gentler but rougher ice sheet margins. This is illustrated with Sentinel-3 data in Figure 3.5 where a derived POCA trajectory and elevation profile are shown together with corresponding waveforms and auxiliary elevation data from REMA (Howat et al. 2019) over all these surface types. We can see relatively sharp and gaussian-type waveforms over the ocean and parts of the (smooth) ice shelf, whereas

the rifted part of the ice shelf and the rough zone of the grounded ice-sheet show more messy waveforms and larger deviations between surface elevation estimates from the two ESA-IPF retrackers (OCOG and UCL ice sheet), as well as from the independent REMA data. We also see that the derived POCA locations are close to nadir over the flatter ice shelf, but deviate by several kilometres across the ice rise and sloping ice sheet margin, with a POCA-to-nadir elevation difference of up to hundred metres according to REMA, and thus beyond the ~60 m range window of Sentinel-3. Track segments with large height difference wrt. REMA (tens of metres) also indicate errors in POCA relocation and hence the surface elevations cannot be trusted locally even if the signal is correctly retracked with a seemingly realistic elevation profile. Also worth noting is that the OCOG retracker is able to retrieve surface elevation over larger parts of the track than the UCL ice-sheet retracker, and that neither of them can capture the steepest part of the ice rise as the onboard closed-loop tracking command (CLTC) is not able to adjust the sampling window to the surface here. This issue could potentially be solved with an open-loop tracking commend (OLTC) informed with an auxiliary DSM like REMA.
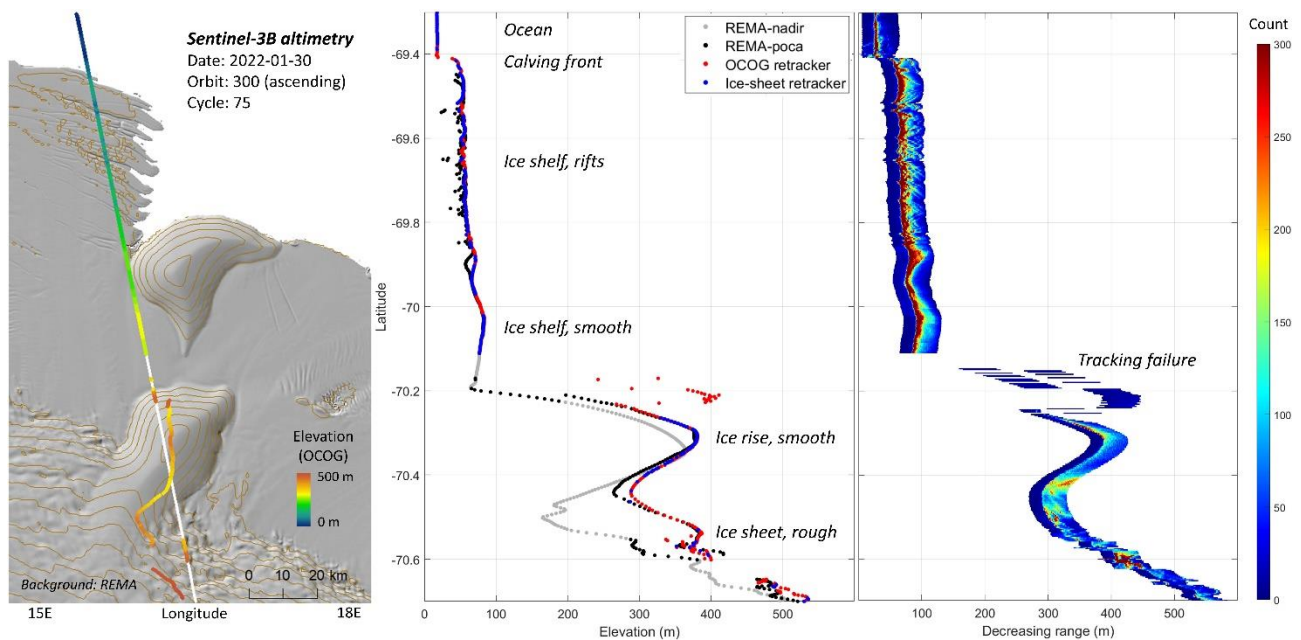
Figure 3.5 Example of a Sentinel-3 track across the coastal margin of the Antarctic ice sheet showing (left) POCA locations and elevations overlied a hillshade from the Reference Elevation Model of Antarctica (REMA, Howat et al. 2019), (middle) along-track elevation profiles as a function of latitude for ESA retrackers and REMA, and (right) corresponding waveforms shifted relative to their tracked range.

Altimetry validation in coastal environments like this is obviously not straight forward, and FRM sites need to consider both the feasibility (coverage of variable POCA locations) and representativity (are the results valid more widely) of the validation activities. For example, a fixed station or in-situ campaign could be centred at a topographic ridge to maximize POCA overlaps, whereas a larger scale airborne campaign could cover more surface types by surveying a grid that would sample POCA tracks more randomly.

3.2 Calibration and validation as a comparison process

Calibration and validation use other independent observations in comparison with measurements by the satellite altimeter to test the altimeter observations. Such comparisons can be made against observations by other satellites or observations from non-satellite methods (e.g., in situ, aircraft, drones, stations). The results of such comparisons can be interpreted in three different ways:

a) A comparison can be used to validate that observation values are within an expected tolerance
b) A comparison can be used to evaluate the uncertainty associated with the satellite observation
c) A comparison can be used to validate independently determined uncertainties.

Traditionally, comparisons have been used for approach (a), i.e., to monitor whether satellite and reference measurements agree within the satellite's requirements. More recently, approach (b) or (c) has been attempted.

In a comparison, it is necessary to consider uncertainties associated with:

(1) The reference / compared observations
(2) The satellite altimeter observations
(3) The comparison process itself

Uncertainties associated with the comparison process itself include uncertainties related to the fact that the reference measurements and the satellite observations may be different, or uncertainties related to processing steps used to make the two measurements more equivalent (for example, by scaling or sampling observations to a common grid, or converting a radar freeboard into a lidar freeboard).

If the uncertainties associated with (1) the reference observations and (3) the comparison process themselves are much smaller than (2) uncertainties associated with the satellite altimeter observations, then the comparison can be used for approach (b), i.e., the comparison can be used to calculate/evaluate the uncertainty associated with the satellite measurements.

A metrological approach to comparisons would follow approach (c). That is, the three types of uncertainty associated with the two measurements (satellite and non-satellite) and the comparison process are independently evaluated and then the comparison is used to validate the uncertainties. It is for this reason, that the FRM needs an uncertainty evaluation independently determined and that we need to consider the uncertainty associated with the comparison process itself.

3.3 Calibration-Validation scheme for a fixed station

Figure 3.4 shows a schematic comparison approach for land ice observations based on in-situ data from a fixed station (for other FRM types, see general discussion in Section 4.1). The satellite nadir point is meant to repeat on a 'theoretical orbit' ground track, however, in practice, due to orbit fluctuations, the actual orbit may differ from this theoretical ground track and may vary from orbit to orbit over scales of hundreds of metres. Actual orbits are shown in the figure as grey dashed lines and are given the count indicator j in the equations given in later sections.
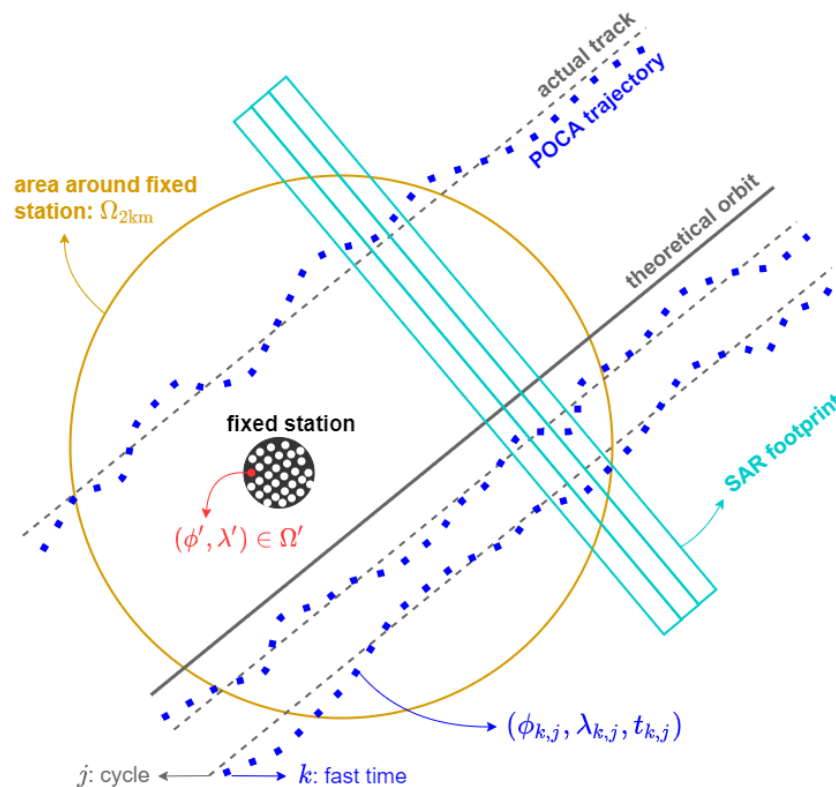
Figure 3.4 Schematic view of a generic scenario for collecting satellite altimetry and in situ data over land ice.

Because the measurement point of closest approach (POCA) is not at the nadir point, as discussed above, the actual observed POCAs are shown by the blue squares. These points will be reasonably close to the nadir ground track for a flat surface and may be further away from the nadir track for a more structured or sloping surface.

The in-situ observations are taken at a selected location near the theoretical nadir track orbit, or where POCA tracks are known to cluster due to surface topography. To account for small scale spatial and temporal surface height fluctuations (discussed further in Section 5.3), measurements are made at multiple locations within a small area (see Section 4.1). This is described in the figure as the fixed station. A GNSS receiver is placed near the in-situ instruments to act as an absolute positioning reference.

A digital surface model (DSM) is needed over the area that covers the possible POCA positions for multiple overpasses along the reference orbit, or orbits – if more than one orbit crosses near the in-situ instruments. This DSM can come from a local measurement campaign (by air or on the surface) and should cover a sufficiently large area to capture deviations in POCA trajectories from the fixed station. For a reasonably flat surface, the area needed can be defined by the potential variability of the ground nadir track, and an area of approximately 2 km diameter is suitable. For sites with significant topographic variability or slope, much larger areas may be needed. The DSM can then be used to transfer POCA heights to the height at the in-situ station or vice versa.

SENTINEL-3 TOPOGRAPHY MISSION ASSESSMENT THROUGH REFERENCE TECHNIQUES (ST3TART)

Ref	NOV-FE-0899-NT-044		
Issue	4	Date	03/04/23
Rev	1	Date	17/04/23
Page	31/67		

4 Requirements for Cal-Val altimetry measurements of land ice

4.1 Types of measurements to include in an FRM suite

There are two fundamental approaches to provide a fiducial reference measurement (FRM) for satellite altimetry observations over land ice. First, there are airborne or in-situ survey campaigns that provide profiles of absolute height along tracks that cross the satellite observations. Such absolute survey campaigns can be performed by moving instruments on the ground (e.g., GNSS on a snow vehicle), or from aircraft (crewed or drone, equipped with lidar, radar or photogrammetry). Airborne sensors can offer complementary information with much finer spatial resolution than spaceborne measurements and wider coverage with respect to in-situ observations. The FRM data from the surveys can be compared directly with the satellite data at the same location, but sometimes requiring spatial interpolation and an assumption that the surface is stable between the times of the FRM survey and the satellite overpasses. The survey data may be resampled to create a gridded comparison either before or after the comparison with the satellite. The time period over which a survey FRM data set is useful for satellite validation will depend on how rapidly the surface is changing and that is highly variable with location. As a rule of thumb, one orbit cycle (27 days) surrounding the FRM survey timing (+/- 13 days) would provide a good compromise between satellite data coverage and FRM stability.
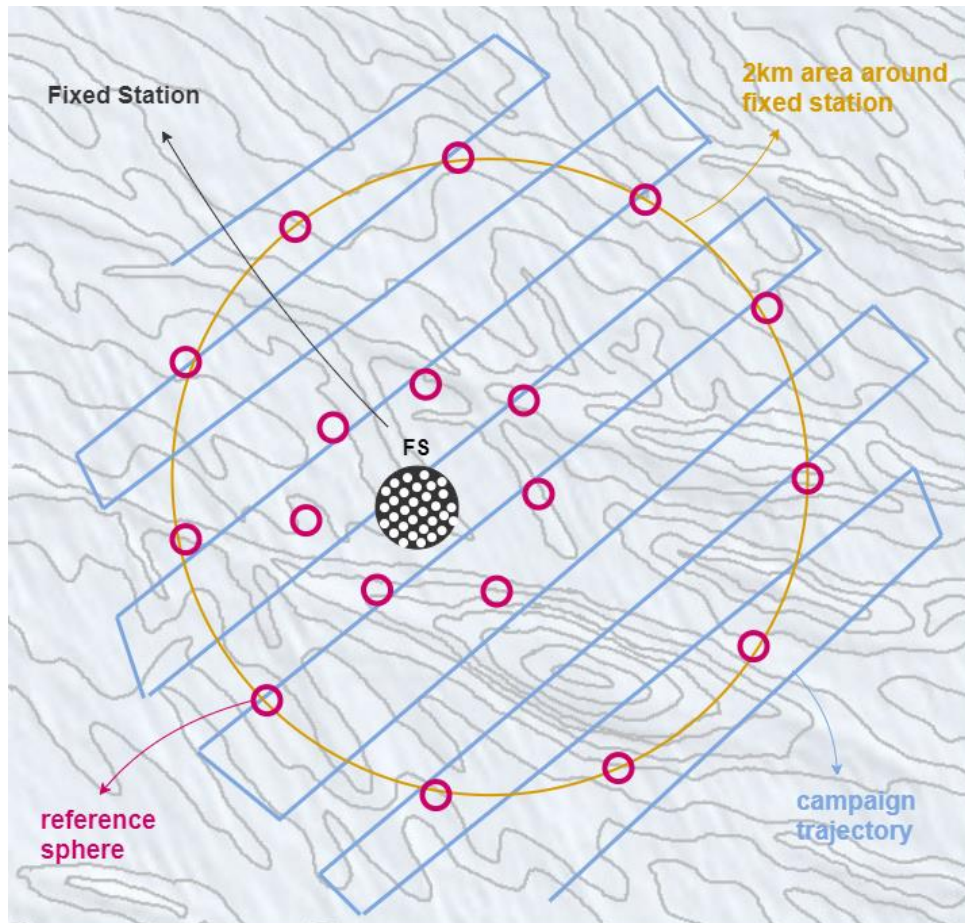
Figure 4.1 Schematic view of a FRM setup for land ice validation consisting of a Fixed Station (FS) and a DSM obtained from a survey campaign

The second approach is to install a fixed ground station that takes pseudo-continuous measurements at a frequency greater than the time between satellite overpasses. To minimize the impact of small-scale surface roughness variations (e.g., sastrugi) and make the FRM surface-elevation time series more robust, the fixed station should ideally consist of several surface-measuring instruments spaced a few metres apart, whose data are checked for errors and averaged to into a time series. A series of satellite observations can be compared to the fixed-station measurements at the time of the satellite overpass. However, because the satellite measurements will differ in location from the fixed station,

especially considering POCA relocation (as discussed in Section 3.1), the fixed station measurements must be supplemented by a high-resolution digital surface model (DSM) over an area of at least 2 km around the fixed station. The high-resolution DSM can be obtained through in-situ or airborne survey campaigns, in the same way as for the absolute survey FRM, but for this application, it will only be used in a relative sense, so high resolution and precision is more important than absolute accuracy which can instead be obtained from the GNSS station. The DSM survey should ideally be repeated annually to account for potential topographic changes, including movement of the ice. Shifting the DSM reference position along with the GNSS-measured movement at the fixed station may seem like a good idea, but topography undulations beyond wind-formed sastrugi and dunes are closely related to bedrock topography and may therefore not advect with the flow like surface features on floating ice shelves and sea ice.

Photogrammetry can provide a DSM with high spatial resolution and high relative accuracy. This method is not directly suitable for absolute validation campaigns but is valuable to supplement a fixed station FRM. There are many ways to obtain images for photogrammetry of snow surfaces: satellite, airplane, UAV (Unmanned Aerial Vehicles) or terrestrial imagery. Based on intercomparison of these photogrammetric platforms (Eberhard et al. 2021), and considering our objectives of resolution, accuracy, size of survey area, and especially the relative convenience of implementation, UAV photogrammetry approach is the main considered option in this document. Lidar systems (Harder et al. 2020) are good alternatives, but generally more complex and expensive to employ. In cases where the surrounding surface can be easily and traversed by snow vehicles, a densely gridded GNSS ground survey is also a good option.

Figure 4.1 shows how a high-resolution DSM based on a UAV photogrammetry can supplement the local measurements at a fixed station. The UAV flies the blue track to obtain a DSM of the site around the fixed station. The indicated reference spheres are large artificial spheres placed around the site to act as guidelines for aligning the photogrammetry images. These reference spheres are arranged both around the fixed station and in a wider ring that surrounds the satellite POCA tracks that the FRM is going to be compared against. If not considering the reference spheres, the figure is also applicable for UAV Lidar and ground-based GNSS surveys where a similar type of survey trajectory would be suitable for generating a DSM around the fixed station.

4.2 FRM quality and other observations

The definition of an FRM (Section 1.3) sets high expectations for FRM-quality measurements. They should be traceable to a community agreed reference (preferably SI) and come with robust uncertainty analysis. These requirements are both important and challenging. It is only with FRM-quality observations that we can perform comparisons to validate satellite observation (and processing) uncertainties (type c comparisons as described in Section 3.2) and therefore support the applications that use sea ice thickness data quantitatively (e.g. in climate modelling and reanalysis).

However, a full FRM-quality analysis of the uncertainties of Cal/Val techniques can be expensive and time-consuming, and for some historical data could be impossible if the full information is not available. Furthermore, the dominant uncertainty in such a Cal/Val process often comes not from the observational uncertainty, but from its representativeness, which means more poorer quality data over a larger representative set of conditions can be preferable to a smaller set of higher quality data over just a few locations / times. Given the limited available information over the challenging polar regions, all data are of value, and therefore we have also looked into cross-validations with the very different ICESat-2 laser altimeter, as summarized in Section 6 with a supplementary technical report [RD9].

In this perspective, it is important to be pragmatic and to consider a Cal/Val process as using information from both FRM-quality and other Cal/Val data sources. However, this distinction should be made clear to those performing Cal/Val analyses, as well as when Cal/Val results are presented.

5 Existing means for Cal-Val activities

5.1 Remote airborne surveys

This section concerns long-range airborne surveys that operate from an airbase, without requiring any field personnel on the ice. Smaller scale in-situ surveys with unmanned aerial vehicles (UAVs) are discussed in Section 5.2.

Airborne surveys over land ice can have multiple purposes:

▲ Reference topography (photogrammetry or lidar with complete coverage, can be an FRM)
▲ Ice thickness and deep stratigraphy (profiling with low-frequency radar, not relevant for FRM)
▲ Snow accumulation and shallow stratigraphy (profiling with high-frequency radar, auxiliary data to an FRM)
▲ Ice thickness change and glacier mass balance (repeated surveys of surface elevation, can be part of an FRM)
▲ Cal-Val of satellite sensor performance over snow and ice (simulations with comparable instruments)

Here, we will not consider the original purpose of airborne surveys, but rather their suitability as past or future FRMs for Sentinel-3 during its mission duration. The two main series of airborne campaigns that have been carried out in polar regions are NASA's Operation Ice Bridge (2009-2019), initiated to bridge the observational gap between ICESat and ICESat-2, and ESA's CryoVEx program (active since 2003 in various forms) that was initiated for cal-val experiments leading up to CryoSat and CryoSat-2, including coordinated in-situ fieldwork on (mainly) Arctic ice caps, the Greenland Ice Sheet and Arctic sea ice. Both airborne programs have had dual lidar/radar systems onboard which has facilitated detailed studies of signal penetration, snow depth and surface mass balance in addition to altimetry validations.

Existing airborne radar and lidar systems for cryosphere applications are described in detail in the TD-2 sea-ice report and are for all practical purposes similar for land-ice use although practical strategies for surveying and implementation in an FRM may vary. Here, we do not go into technical details but rather focus on the relevance for land-ice validation.

A summary table of existing and planned campaign data suitable for Sentinel-3 FRM is provided in the roadmap (TD-6).

5.1.1 Radar

Detailed surface elevation and thickness surveys have been carried out since the 1990s over polar land ice (e.g., Bindschadler and Vornberger, 1994). At that time the uncertainty was poor, several meters for surface elevation (Retzlaff and others 1993). Since then, many airborne operations have been conducted joining radar and lidar measurements to retrieve ice sheet surface elevation, as well as firn/ice layering and thickness.
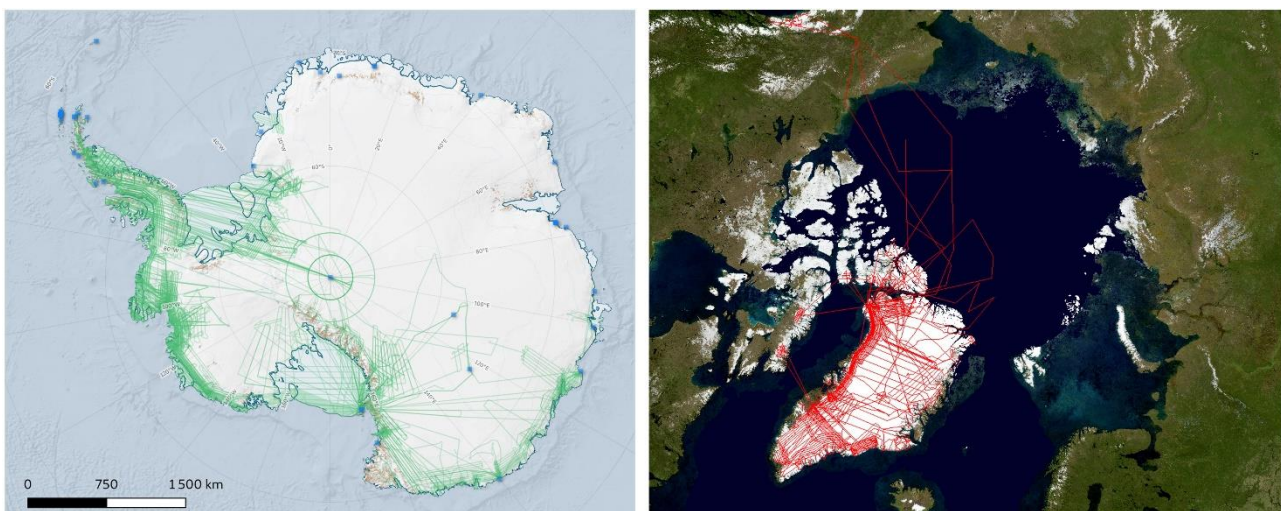

Figure 5.1 - *Operation Ice Bridge tracks in Antarctica for all years (green lines, left) and for a selected year 2014 in the Arctic (red lines, right).*

The most important operation in terms of data coverage has been NASA's decade-long Operation IceBridge. It was an airborne mission to survey land ice and sea ice across Antarctica, Greenland, the Arctic archipelagos, Alaska and Iceland. IceBridge conducted 968 science flights, of which 42% were repeat surveys of land ice (MacGregor et al., 2020). Figure 5.1 show examples of IceBridge flightlines over Antarctica for all years and over the Arctic for 2014 as a random year.

Operation IceBridge used different radar instruments operated by the Centre for the Remote Sensing of Ice Sheets (CReSIS) at the University of Kansas, among them the CReSIS Ku band radar altimeter. The data from this radar can be used together with lidar data to determine the thickness of snow over sea ice. The Multichannel Coherent Radar Depth Sounder (MCoRDS), (Rodríguez-Morales and others, 2014) operating with a 180 MHz – 210 MHz frequency range with multiple receivers and an adjustable bandwidth up to 30 MHz, was also used to retrieve ice thickness with a post-processing vertical resolution of 4.5 m in ice an along-track sampling of 30 m (Leuschen and others, 2000). The CReSIS ultra-wideband FMCW radars include the snow radar (2 GHz – 8 GHz), the Ku-band Radar (12 GHz – 18 GHz), and more recently the Snow/Ku Radar covering the entire bandwidth of those first two systems (2 GHz – 18 GHz). A Ka-band (32 GHz – 38 GHz) radar was also operated during one campaign (2015 Arctic spring). The Snow Radar, was used to get the stratigraphy (isochrones) in the upper 20 m – 30 m of the ice sheet with fine vertical resolution (e.g., Medley et al., 2014); while the Ku-radar was used to follow reflection horizons below the surface related to different scattering characteristics of the firn including melt features (Kuipers Munneke et al., 2017). An Accumulation Radar, with an ultra-wideband stepped-frequency chirped pulse radar system operating between 600 MHz and 900 MHz, was used to image horizons in the upper 300 m of the ice sheets (e.g., Medley et al., 2014). Finally, when possible, IceBridge surveys included near-coincident satellite underflights (mostly CryoSat-2), coordinated flights with ESA's CryoSat Validation Experiment (CryoVEx) airborne campaigns.

As a part of the CryoSat Validation Experiments (CryoVEx) carried out in the Arctic and Antarctic since 2003, ESA used a CryoSat-like Ku-band radar altimeter known as the Airborne SAR/Interferometric Radar Altimeter System (ASIRAS). ASIRAS was operated together with airborne laser scanner (ALS) and vertical photography for joint data acquisitions and analyses over selected sea ice and land ice validation sites and CryoSat ground tracks. ASIRAS was used to track the annual snow accumulation layers of the upper firn column at regional scales over the ice sheets of Greenland (Hawley et al., 2006; Overly et al., 2016) and Antarctica (Kowalewski et al., 2021), as well as selected Arctic ice caps, mainly Devon Ice Cap in Arctic Canada and Austfonna on Svalbard (Hawley et al. 2013). ASIRAS operates at a carrier frequency of 13.5 GHz and a bandwidth of 1 GHz (Mavrocordatos et al., 2004). CryoVEx/ASIRAS has been followed by CryoVEx/KAREN, which is a Ka-band radar with central frequency 34.525 GHz, operated in several airborne campaigns over land ice and sea ice since 2016. More details about these radar altimeters can be found in the associated TD-2 sea-ice report.

It should be noted that the FRM considerations of this report has been centred around validation of the intended Sentinel-3 end-product over land ice: surface elevations. Although airborne radar altimeters can be extremely useful for investigating signal backscattering properties at comparable frequencies as Sentinel-3 SRAL, they share the same issues of penetration and potential biases with respect to the actual surface. Hence, airborne radar altimeters should be supplemented by sensors that confidently measures the actual surface, of which lidar is the preferred choice.

5.1.2 Lidar

NASA's IceBridge program (Blair and Hofton, 2012; Krabill, 2013) has used three different lidar systems: the Airborne Topographic Mapper (ATM, Martin et al., 2012), the Land, Vegetation, and Ice Sensor (LVIS, Hofton et al., 2008), and the scanning photon-counting lidar (PCL) coupled to a legacy nadir-pointing laser profiler (LAS), making the Airborne Laser Altimeter with Mapping Optics (ALAMO, Young et al., 2017). ALAMO was developed at the University of Texas and was supported by the ICECAP (International Collaborative Exploration of the Cryosphere by Airborne Profiling) consortium. In particular, ICECAP collected surface-elevation data over several years in otherwise sparsely surveyed areas in East Antarctica. The ATM lidar system also has a history before IceBridge and has been operated since the early 1990s (Krabill et al. 1995). Since then, ATM surveys have been carried out almost every year over the Greenland Ice Sheet, and also over Arctic ice caps (Canada, Svalbard and Iceland) and the Antarctic Ice Sheet in selected years. IceBridge flightlines with lidar are similar to those with radar (Fig. 5.1). Although the data coverage looks very dense at a continent-wide scale, the actual lidar data are limited to relatively narrow lidar swaths of a few hundred meters across-track depending on flight altitude.

ESA's CryoVEx program has used the Airborne Laser Scanner (ALS) system operated by the Danish Technical University or the Alfred Wegner Institute in various survey campaigns, usually carried out in spring with targets both over Arctic land ice and sea ice. ALS has a sub-metre horizontal resolution and a swath width that is roughly equal to the flight height above the ground, from a few hundred meters up to a kilometre depending on the lidar model used. The ALS vertical accuracy is stated to be around 10 cm [RD11], including errors from the GNSS positioning and the inertial navigation system (INS) for attitude determination. The high accuracy and minimal penetration make such data ideal for absolute validation of surface-elevation products from satellite altimetry (e.g., Sørensen et al., 2018).

In populated areas, airborne lidar has become the new standard for topographic mapping and many commercial operators exist within Europe. For example, almost the entire Scandinavia has been scanned with lidar during the last decade, including mountain glaciers. This is not yet the case for any of the Arctic archipelagos like Svalbard, but could be so in the future. If this is to happen it could provide unique opportunities for satellite altimetry cal-val over large areas with a wide range of glacier types and surface topographies.

5.1.3   Photogrammetry

Airborne photogrammetry has for long been the preferred technique for detailed topographic mapping and has been extremely useful for quantification of long-term glacier changes since the first stereo-photo campaigns (often oblique) in the early 1900s (e.g., Geyman et al. 2022). There has been a number of historical photogrammetry campaigns for first-time mapping along the coasts of Greenland and Antarctica, but in recent decades, such dedicated large-scale topographic surveys have been limited to populated areas and some of the Arctic archipelagos, in particular Svalbard where the entire archipelago was mapped with sun-metre stereo photos during the period 2008-2012. Such data are generally not sufficiently accurate for absolute validation of satellite altimetry, but their high relative accuracy can make them suitable as high-resolution DSMs for spatial transfer of FRM and satellite observations at different locations. However, the limiting factor of these high-quality DSMs is that they get rapidly outdated over changing land-ice surfaces, and that they are rarely updated in polar regions, just like for complete area surveys with lidar.

5.2   In-situ surveys

In-situ surveys of surface elevation are done with kinematic GNSS with the antenna mounted on a snow vehicle/sledge or by a UAV using photogrammetry or lidar. In an FRM context for Sentinel-3, these surveys could be targeted at high-resolution DSM generation around a fixed station (Section 3.3) or larger scale surveys aimed at direct validation of multiple Sentinel-3 tracks over a given period (e.g. one S3 repeat-cycle) where the surface can be assumed to be stable. For the first purpose, it is sufficient to have a DSM with good relative accuracy as it is used to transfer observations between the satellite overpasses and the fixed station, whereas for the second purpose the survey also needs to have a very high absolute accuracy, which can be difficult for UAVs that have limited instrument capacity and not always possibility to carry high-level geodetic GNSS and Inertial Navigation (INS) systems, unlike manned aircrafts.

5.2.1   Surface GNSS

Kinematic GNSS data suitable for altimetry cal-val are typically collected in localized observational campaigns or over longer transects as part of logistical traverse routes, mainly in East Antarctica. The GNSS receiver can be mounted directly on the snow vehicle or on a sledge that is pulled behind (Fig. 5.2). Surface receiver position is derived from absolute precise-point-positioning (PPP, Section 5.2.1.1), involving delayed post-processing when precise correction data become available, or relative to a fixed base station in a differential manner (Section 5.2.1.2), in real time with radio link or post-processing without. The observational frequency should be high enough to get a dense observational sampling along the survey track, e.g. 1-5 Hz. GNSS observations should rely on dual-frequency carrier phase, not just code, to achieve sufficient accuracy. GLONASS and GALILEO capability is also recommended for better satellite geometry than GPS alone. Surface elevations are derived from the GNSS data by subtracting the antenna height over the surface which is measured by a tape measure. Compression of surface snow by the vehicle should be accounted for in the measurement of antenna height, and this can vary with time and location.

Surface GNSS surveys can be targeted on DSM generation around a fixed station, involving a grid-line survey with interpolation between tracks, or on specific Sentinel-3 ground tracks by grid-line surveying around an ascending-descending crossover location or where POCA tracks tend to cluster along topographic ridges such as ice divides. If the location of POCA tracks can be precisely predicted, or the survey carried out shortly after a satellite overpass, then the

navigational flexibility of a snow vehicle makes it possible to follow an undulating POCA track exactly on the ground for direct absolute validation. This can be an efficient way to achieve many comparison points for individual tracks.

Figure 5.2 Example of geodetic GNSS mounted on a snowmobile sledge (left) and a tracked snow vehicle (right) for kinematic survey.

5.2.1.1   Precise Point Positioning (PPP)

This approach relies on carrier phase measurements without any base station. The first-order ionospheric delay is removed from the combination of several dual-frequency GNSS measurements. GNSS satellite orbit and clock errors corrections are derived from a network of global reference stations and can be directly delivered by geostationary satellites, but it is preferred to use more precise post-processed products that become available with a few weeks delay. The wet tropospheric delay is typically corrected with a meteorological model, whereas as the wet part can be treated as an estimate of the Extended Kalman Filters (the core of the PPP algorithm) as for carrier phase ambiguities. The methodology then relies on iteratively minimising the noise along with additional GNSS measurements until convergence is reached under a given threshold. This technic precludes real time approaches as it relies on the delivery of corrections from the global GNSS network and requires subsequent computational time. Typically, an uncertainty within 10 cm can be achieved depending on the quality of corrections. The main advantage of the PPP method is that it does not require any base station which is difficult operate in remote polar areas.

Below, the significant sources of GNSS uncertainty are listed, along with a description of how they are dealt with. Most of these sources of uncertainty are common for airborne (and satellite) observational systems as they also rely on GNSS for positioning of other sensors like radar, lidar and optical cameras.

Ionospheric delay. Partial ionisation of the ionosphere by UV radiation produces free electrons whose variable content TEC (Tropospheric Electros Content) delays the wave travel time leading to a range error of the order of 5 m. Fortunately, this delay is frequency dependent and proper combination of the generally available L1 and L2 frequencies allows first order assessment and subsequent removal of this artefact. The residual uncertainty associated with the correction is sufficiently small for our application and smaller than other uncertainty sources.

Tropospheric delay. Meteorological conditions of the troposphere (mainly the water content) affect the wave speed and, as with the tropospheric crossing, induce a corresponding delay (of the order of 0.5 m) which is unfortunately not frequency dependent. One method to correct for this delay uses pressure data from atmospheric reanalysis. Such methods work well in well-studied areas where the reanalyses are supported by many observations. It is more difficult to assess the resultant uncertainty on the correction for very remote places like Antarctica where constraining meteorological observations are very sparse. The resulting uncertainty is not well known in Antarctica.

Satellites ephemeris and clock errors. Despite a very strict control of the so-called 'control segment' allowing accurate updates of both satellite ephemeris and clocks, these last two source error terms potentially degrade range solution by

respectively some 2.5 and 2 m. Post-processed products using the global network of GNSS receivers are able to improve these corrections substantially and should therefore be applied before altimetry validation.

Receiver noise. Additional range errors arise from the receiver hardware and software noise. However, these issues reduce along with the quality (and cost) of the receiver as well as with the fast-evolving algorithms.

Multipath. An error term that can be significant and which depends on the measurement surroundings. It represents all means by which the satellite signal does not directly reach the receiver owing to nearby potential intermediary spurious reflection surfaces (mountain, vehicle, etc.). The best remedy lies in ensuring a clear horizon to the GNSS antenna and avoiding any nearby obstacles. A minimum threshold of satellite elevation angle can also help, especially relevant in polar areas where satellites can be seen across the poles but often at low angles. Also, antennas equipped with ground plates can significantly filter out multipath incoming rays.

5.2.1.2   Differential GNSS

The major interest of the differential approach lies in the fact that most errors are spatially correlated and can therefore be eliminated provided the distance between the base station and the rover (the so-called baseline) remains low, of the order of this correlation distance. The method states that the range difference between the actual base station position (assumed to be perfectly known) and the one it computes represents the various error contributions. This error term can therefore be transposed to the rover and incorporated in its position calculation.

The requirement of a base station locations with proper positioning can be tedious because of the required long-distance links to existing reference networks. Moreover, the most stringent condition comes from the fact that these reference stations should remain fixed in time which excludes all icy spots and requires rocky ground reference stations (generally located on the coast). Traditional baselines of the order of a few tens of km therefore strongly forbid the use of this methodology over the inland parts of the ice sheets. A way to circumvent this last drawback is to retrieve corrections from dedicated systems called satellite-based augmentation systems (SBAS). Unfortunately, these systems have only been developed for populated areas (WAAS for the USA, EGNOS for Europe) and exclude the polar regions.
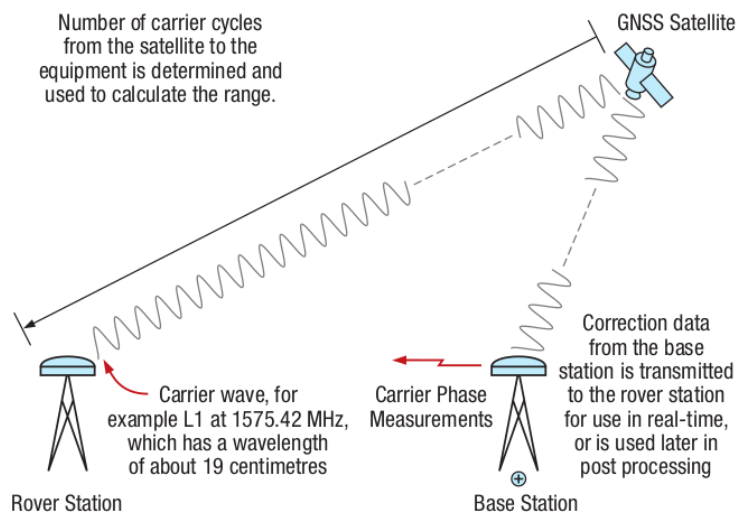
Figure 5.3 - Basic principles of the RTK position. L1 wavelength being 19 cm, the phase reading resolution of a couple of degrees maps into theoretical intrinsic range accuracy of less than 1 cm.

Differential GNSS on the carrier phase can be carried out as Real Time Kinematic (RTK) or Post Processing Kinematic (PPK). The RTK approach relies on carrier-based ranging and baseline computations within a base station network, from one to several reference points. For absolute range calculation, the phase ambiguity must be resolved. The ranging distance between the satellite and the receiver equals the number *N* of carrier cycles times the wavelength (plus the final non-integer final cycle). However, the reading at the receiver is only a 0° - 360° angle with a priori no access to the total number of cycles and hence to the actual ranging (only time dependent changes in distance can be inferred as long as the receiver 'locks' onto the satellite). Solving this ambiguity (a posteriori determination of *N*) is the challenge of this

approach and is carried out by operating single, double and triple differences of at least 2 receivers locked onto at least 4 satellites. For real-time applications, this therefore requires communication links between roving and reference receivers which are generally done with real-time links with radio-modem devices. In case the real-time link cannot be achieved, the PPK approach can be used where both receivers record data independently and are linked after the survey with an appropriate PPK software. Baseline correlated errors are then dealt with similarly as for RTK. The accuracy of either approach under good conditions is on the cm-level.

Some examples of GNSS surveys and related uncertainties are presented in Table 5.3 for Antarctica.

Table 5.1- Examples of various GNSS survey types caried out in Antarctica

Method/survey	Place/date	obtained Accuracy	Comments	Reference
PPP Permanent monitoring of the surface velocity field of acoastal outlet glacier Permanent self-autonomous network of 8 GNSS beacons. Trimble Net-R9 dual frequency geodeic receivers	Astrolabe coastal Glacier East Ant. 139°55' E 66° 44' S 2013-today (supposedly continuous)	10 cm depending on time-evolving constellation quality	Harsh environment (especially windy) with time-limited (summers) and difficult access (crevasses)	Ducasse et al., 2015
RTK Tidal surface displacements over the floating part of a coastal glacier in order to constrain its grounding line position Leica dual freq. Geodetic receivers GX1230	Astrolabe coastal Glacier East Ant. 139°55' E 66° 44' S 16-18 jan 2011	5 cm which allowed to infer small (attenuated) tidal displacements Baseline length of less than 10 km	Required height measurements of exactly the same locations at 12-hour interval. Surface height precision hampered by mean time glacier displacement when the slope is significant (up to 3 cm extra error term)	Le Meur et al. 2014)
RTK Roughness measurements over limited areas (100-200 m²) along scientific traverses Leica dual freq. Geodetic receivers GX1230	Along the EAIIST scientific traverse over the East Antarctic Plateau Dec 2018- Jan 2019	2 cm to 3 cm thanks to a very short baseline (a few tens of m). Most of the error due to antenna positioning over targets	No real reference station used here in the sense that the base station was simply dropped anywhere and the resulting 'navigated solution' used for sending corrections. Although not geographically accurately positioned all measured points where all consistent to one another to within the 2 cm to 3 cm accuracy owing to the correctness of the resulting (short) baselines	Larue et al., 2021
PPK Ice-surface elevation profiles over the east Antarctic plateau in order to validate altimetric data from satellite radar and laser Geodetic dual-frequency GNSS receivers	Around Vostok Lake (106.8 ◦ E, 78.5 ◦) Sand along vehicle traverses between Vostok and Mirny stations then bet. Vostok and Progress 2001-2015	4-9 cm allowing for assessing time-dependent changes of potentially stable surfaces. Accuracy potentially seriously hampered by vehicle sinking and tilt over soft and bumpy areas but partially corrected for.	Very long baseline (up to 800 km) ranges with reference to Casey and Davis stations that questions the validity of the method. Many phase slips that required tedious manual checking and removal of a significant number of observations Post-processing with the help of the Bernese GNSS Software 5.2 (Dach et al., 2015)	Schroeder et al., 2017

5.2.2   UAV photogrammetry

In-situ surveys of surface topography can also be carried out with unmanned aerial vehicles (UAV), typically using photogrammetric techniques to generate a high-resolution Digital Surface Model (DSM). The development of surveying techniques to obtain DSM by SfM (Surface from Motion) analysis of UAV images has been very important in recent years. Motivated by applications in hydrology, there are many applications for snow height measurements in mountainous areas, based on the subtraction of DSM with and without snow (Revuelto et al., 2021, Adams et al. 2018,

Avanzi et al. 2018, Buhler et al. 2016). Despite the lack of studies on Antarctic snow, it seems reasonable to apply the protocols and uncertainty estimates proposed in the literature of snow surfaces in mountainous areas.

Due to logistical constraints, DSMs can only be obtained during the summer months in Antarctica. The use of an UAV Lidar system allows the mapping of snow surfaces with final DSMs of similar uncertainties as those obtained by UAV photogrammetry SfM (Harder et al., 2020), and UAVs have already been tested during summer 2021-22 on the Astrolabe glacier near Dumont d'Urville station (unpublished).

Most UAV photogrammetry snow studies use Agisoft Metashape software (SfM workflow) to generate DEM from a series of images (Goetz et al. 2019, Adams et al. 2018). A specific workflow described in the USGS workflow (Over et al., 2021) and has already been adapted to the specificities of snow surfaces (high reflectance and low contrast images).

The uncertainty of DSMs form UAV photogrammetry depends on the type of surface. Several studies have investigated uncertainties over different surfaces by comparison to in-situ measurements. Table 5.2 summarises the uncertainty levels published in the literature. These uncertainties are determined with respect to absolute heights. From these and despite the diversity of assessment methods, the expected accuracy of the DSM measured by the RMSE, are expected in the range of 10 cm to 20 cm.

Table 5.2. Absolute uncertainty of UAV photogrammetry (SfM) on snow surfaces

Location / field type	Variable	Uncertainty	Ref
Mountain snow Spanish Pyrenees	Mean RMSE (DSM)	7 cm – 24.6 cm / TLS	Revuelto et al. 2021
Mountain snow Spanish Pyrenees	Mean RMSE (snow depth)	22 cm / TLS	Revuelto et al. 2021
Alpine snow	Mean RMSE (DSM)	7.5 cm / GCP	Harder et al. 2016
Glacier w/wo snow Swiss Alp	Mean RMSE (DSM)	10 cm – 25 cm /GCP	Gindraux et al. 2017
Alpine snow in alps Switzerland	Mean RMSE (snow depth)	21 cm /manual measurements	Eberhard et al. 2021
Mountain snow Tasmania	Mean RMSE (DSM)	6 cm /GCP	VanderJagt et al. 2015
Mountain snow Tasmania	Mean RMSE (snow depth)	9.6 cm – 18.4 cm /manual measurements	VanderJagt et al. 2015
Alpine snow - plateau area - Italy	Mean RMSE (snow depth)	14.3 cm /manual measurements	De Michele et al. 2016
Alpine snow in alps Switzerland	Mean RMSE (snow depth)	7 cm – 30 cm /manual measurements	Buhler et al. 2016
Alpine snow - slope scale - Austria	1 sigma (DSM)	< 25 cm / TLS	Adams et al. 2018
Polar snow Alaska	sigma 95% confidence (DSM)	+- 30 cm	Nolan et al. 2015 *

For the FRM, however, the DSM will only be used relatively – to determine the relative height difference of the fixed station to the satellite POCA point. Therefore, of more interest is the uncertainty associated with the relative DSM. Table 5.3 provides indications of the repeatability of measurements at the different sites studied.

Table 5.3 - Precision (repeatability) of UAV photogrammetry (SfM) on snow surfaces

Location / field type	Variable (sigma – std. dev.)	Relative Uncertainty / Repeatability	Ref
Alpine snow - slope scale - Austria	1 sigma - 12 experiments (snow depth)	4 cm small stable area < 33 cm whole study	Adams et al. 2018
Alpine snow	1 sigma - 13 experiments (DSM)	2 cm	Goetz and Brenning 2019
Polar snow Alaska	sigma 95% confidence - 2 or 6 experiments ((DSM)	3 cm from 2 experiments without changes in surface elevation 8 cm for all 5 experiments	Nolan et al. 2015 *

*Nolan et al. study is based on airborne photogrammetry SfM with manned aircraft.

Some studies (Nolan et al. 2015, Goetz et al. 2019, Adams et al. 2018, Avanzi et al. 2018) propose estimates of the precision (repeatability) of DSM on snow surfaces between 6 to 20 cm.

5.2.3   UAV Lidar systems

Unmanned aerial vehicle (UAV)-based equipped with light detection and ranging (LiDAR) technologies is another technology applied to retrieve glacier surface topography and produce three-dimensional digital surface model (DSM) (Bhardwaj et al., 2016). This method has been initially presented by Crockert et al., (2012) who used sensors designated as CULPIS (The Colorado University LIDAR Profilometer and Imaging System). Associated with synchronous GNSS in PPP mode, the LiDAR could allow a centimetre-level accuracy (Mertes et al., 2017), even though a lower accuracy than 10 cm was obtained by Crockert et al. (2012) in Greenland, Svalbard, and over the Southern Ocean (near McMurdo station). Currently, only rare works have been achieved on the ice topography in polar regions. This technique has been used to retrieve mass changes and such as calving volume calculation (Köhler et al., 2019), and recently used to characterize the spatio-temporal evolution of the Dalk Glacier surface. In the latter study, they note that compared with photogrammetry, LiDAR systems present a better capability to characterize ice morphological details, whereas photogrammetry is cheaper and more time-efficient for large-scale survey (Zhang et al., 2020). In this study, they used a Velodyne VLP-16 model mounted on a DJI Matrice 600 Pro UAV.

5.2.4   Auxiliary observations – snow radar

Ground penetrating radars (GPR) have been largely used over ice sheets since their first use over ice in 1964 (Walford, 1964). Glaciological applications are large, including the retrieval of glacier or ice sheet thickness (e.g. Welch et al., 1998), basal conditions (e.g. Winebrenner et al., 2003), liquid water content (e.g. Arcone, 1996), and internal structure (e.g. Pälli et al, 2003). Ground penetrating radar is generally used to map the ice thickness thereby leading to the elevation of the bedrock (using wavelengths around 1-50 MHz) and / or the spatial variation of snow accumulation (using wavelengths of 50MHz-2.5GHz).

In Antarctica, measurements were generally done along traverses, more particularly in the framework of international SCAR project ITASE (International Trans-Antarctic Scientific Expedition) (Mayewski et al. , 2005). This was the case from Troll station to South Pole (Anschütz et al., 2009), from the Swedish Wasa station to the Japanese Syowa station (Fujita et al., 2011) and along the French traverse to Dome C (Verfaillie et al., 2012), between Concordia and Vostok Stations (Lemeur   et al., 2018), from Ross Sea to Talos Dome (Frezzotti et al., 2005), or along Pine Island Glacier catchment in West Antarctica (Konrad et al. 2019).

The principle of ground penetrating radar (GPR) consists of measuring the two-way travel times of parts of an electromagnetic pulse emitted toward the ground that are reflected on discontinuities within the observed medium. Discontinuities caused by physico-chemical transitions and resulting changes in the complex dielectric constant ϵ^* of the medium are responsible for partial reflections of pulses emitted from points at the surface. For the snowpack, and assuming the complex representation of the electromagnetic wave, the permittivity, which varies with density, induces changes on the real part, whereas the conductivity of the medium acts on the imaginary part of the electromagnetic pulse (Eisen et al., 2008). As a result, within the first few hundred meters of the snowpack, changes in firn and ice density and acidity from volcanic deposits create radar reflectors.

The internal reflection horizons inherit the expected synchronism of the responsible physical/chemical processes t. Then, assuming isochronicity, the spatial variations of the depth of the internal reflection horizons reflect changes in snow accumulation. This has only been demonstrated by connecting radar horizons with accurately dated layers in ice cores at different drilling sites (Eisen et al. , 2004; Frezzotti et al. , 2005) or by comparing accumulation rates assessed by a GPR analysis to direct measurements on stake lines (Vaughan et al., 2004). However, the physical processes leading to these internal reflection horizons still remain poorly known.

The methodology implies transforming the two-way travel times of the reflected waves into actual depths, which is only possible if the wave velocity within the medium is known. To convert travel times into depths a vertical velocity profile for the radar wave is assumed. Generally, this is done by considering only the effect of the firn density on the electromagnetic wave velocity (as a result of dry and clean snow over the Antarctic Plateau) (e.g., Verfaillie et al. , 2012). Finally, determination of the snow density is fully necessary to convert the thickness of the snow layer between the internal reflection horizon and the surface, (or between two separated horizons) into a snow accumulation in water equivalent (e.g., Eisen et al., 2008).

5.3 In-situ measurement stations

For S3 validation, the measurand (FRM) is the surface height obtained with time series of static AWSs. At one point, it could be obtained thanks to two complementary sensors: a GNSS for the absolute position of the station, and a one or more surface height rangers for the relative position of the surface wrt. to the station (e.g., Ryan et al., 2008). The measurand is derived as follows (Figure):

$$XYZ_{surf} = XYZ_{GNSS} - R_{ULS} - XYZ_{offset} \qquad \text{(Eq. 6.1)}$$

Where XYZ_{GNSS} is the georeferenced position of the antenna, XYZ_{offset} is the height offset between the GNSS position and the sonic ranger (supposed fixed if the station is not tilting with time), and R_{ULS} is the range under the level of the sonic gauge until reaching the snow surface.

In Section 7.4.1, this same equation is rewritten as
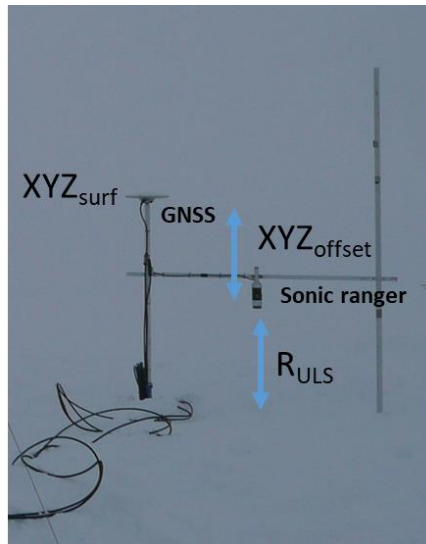
Figure 5.4 - Automatic weather station showing GNSS and ultrasonic rangers.

Hereafter we detail how these differences are obtained in the field with sonic and optical (laser) gauges and GNSS, and how the uncertainty of measurements is retrieved. XYZ_{offset} is the vertical height between the GNSS sensor and the ultrasonic ranger. After installation, this variable only depends on the station mast tilt and deformation.

RULS is typically mounted as a part of Automatic Weather Stations with sensors designed to assess the snow depth on the ground. Typically, the surface range is obtained by timed ultrasonic wave transmission and reflection, but laser

snow-depth sensors also exist. Unlike ultrasonic gauges, the optical sensors are more independent of temperature changes and wind speed, but they are still significantly impacted by precipitation (e.g., SPICE final report, 2018) and the footprint is rarely more than centimetre wide. They are also more expensive and consume much more power. The main interest of these sensors is to collect high frequency continuous data, allowing access to the temporal variability of elevation. The sonic rangers are also known to be robust and very simple and suitable for use in the field (e.g., Eisen et al. 2008).

Several ultrasonic snow depth sensors exist in the industry, as the Campbell Scientific SR50A, the Felix Technologies SL300, and the Sommer USH-8 whereas optical sensor also exist as the Lufft/Jenoptik SHM30 (now available as SHM31), Dimetix FLS-CH 10, or the Campbell-Scientific SDMS-40 Multipoint Scanning Snowfall Sensor. Except for the SDMS-40, all these snow-depth instruments were tested at different sites in various regions of the globe during WMO Solid Precipitation Intercomparison Experiment (SPICE) in order to assess their reliability for snow depth measurements. All sensors were demonstrated to present a similar mean bias (of about 2 cm) and RMSE (1.5-1.7 cm) with observed snow accumulation values. The SL300 snow-depth sensor sometimes had reliability issues that may be site dependent (SPICE final report, 2018).

5.3.1 Absolute reference height - GNSS

The GNSS technique is described in Section 2.2.1 and the same principles apply for a fixed station, typically using precise-point-positioning (PPP) techniques at a delay when precise correction data from the global GNSS network is available.

The purpose of recording GNSS data on a land-ice station is typically:

▲ To track ice movements, both horizontally and vertically (submergence/emergence and hydrological uplift)
▲ To derive snow accumulation variability by GNSS reflectometry
▲ To derive surface elevation changes by combining absolute GNSS positioning with relative surface ranging

For altimetry cal-val, it is the latter purpose that is crucial, and we need to consider both the GNSS instrument itself and the instrument that tracks the snow level relative to the GNSS, typically one or more ultrasonic rangers installed as a part of an automatic weather station (AWS).

5.3.2 Relative surface height – ultrasonic rangers

Ultrasonic gauges are simple to set up in the field, and have been installed in Antarctica over more than 20 years (e.g., Eisen et al., 2008). Several gauges have been installed under various climatic conditions in different regions of Antarctica, as for instance in Princess Elizabeth Land (e.g., Dahe et al., 2004), Adelie Land (e.g., (Favier et al. 2011; Amory et al. 2016)); on the East Antarctica Plateau, at Dome A, Dome F, Vostok or Dome C (e.g. (Libois et al. 2014)), in Dronning Maud Land (e.g., Reijmer et al., 2003), in West Antarctica West Antarctica (van Lipzig et al., 2004), on the Ross Ice Shelf (e.g., Knuth et al., 2010; Cohen and Dean, 2013; Seefledt et al., 2021), in the dry valley (Fountain et al. 2010) and Taylor Valley (Myers et al. 2020), etc. ultrasonic Depth gauges are becoming a standard addition on University of Wisconsin's AWS (Lazzara et al. 2012), which is the main AWS network in Antarctica (http://amrc.ssec.wisc.edu/aws/index.html).

Ultrasonic gauges measure the two-way time of flight of ultrasonic pulses between the surface and an emitter, and register the time of the first echo (t_{first_echo}), similar in principle to the altimeters. In particular, the measurements must take into account the dependence of sound speed with air temperature along the travel path in order to convert t_{first_echo} into a distance (RULS). Temperature should be measured in principle at all levels between the sensor and the surface (due to the strong temperature gradient in stable conditions as in winter). In practice at least one shielded and naturally or mechanically ventilated thermometer approximately halfway between the ultrasonic height ranger and the surface is necessary (Eisen et al. 2008). The sensors should be installed vertically, but the ultrasonic beam allows the sensor angle to be spread over 22° from the vertical (Eisen et al., 2008). The addition of an inclinometer to the sensor can be proposed to verify that the tilt of the station does not cause divergent and irrelevant data. The sensors operate over a height from 0.5 m to 10 m, and sensors should be set up at a significant height above the surface when snow is accumulating. However, the absolute accuracy of the sensor decreases when the sensor is installed high above the surface (e.g. SPICE final report, 2018). The footprint also changes in size with the installation height.

The most commonly used sensor in Antarctica is the Campbell Scientific SR50A. The accuracy given by the manufacturer is the maximum value between ± 1 cm and 0.4% of the measured height, if the impact of air temperature is accurately corrected. In the SPICE final report (2018), the authors found a mean bias of 2.14 cm instead and a root mean square error of 1.64 cm. This maximum accuracy of 1 cm is thus insufficient to identify each snowfall event as occurring on the

Antarctica Plateau, i.e., when solid precipitation frequently results in less than 1 cm accumulation (Reijmer et al., 2003, Picard et al. 2019). SR50A, as all ultrasonic gauges, are also susceptible to failure, and have a membrane that can be destroyed, obstructed by snow/ice, or buried. The echo can also be erroneous if an obstacle or dense medium is encountered beneath the transmitter. This is the case during solid precipitation and, over snow-covered surfaces, when high wind speed causes snow drift in the near-surface air layer (e.g., Li and Pomeroy 1997; Mann et al., 2000, Amory et al. 2021, SPICE final report, 2018). Snow drift can also lead to the formation of cornices or excavations downwind of obstacles. The sensors must therefore be installed far from any obstacle to avoid any spurious accumulation/erosion (Eisen et al., 2008). Finally, the sensor only informs on the elevation range below the sensor and does not inform on the spatial variability of surface changes. This is critical in Antarctica where large surface roughness features are encountered (e.g. Sastrugis, dunes). Thus, the deployment of multiple sensors is necessary to better document surface changes (e.g., Eisen et al., 2008; Picard et al. 2019).

5.3.2.1   Impact of temperature

The impact of air temperature (T_{air}) on the speed of sound is corrected using measurements done generally between the sensor and the surface. However, larger biases are generally found for low temperatures when comparisons are made with manual observations. The sensor has been shown to operate down to -70°C (Van den Broeke et al. 2004, Libois et al., 2014), but these values are well below the specified operating range of the sensors (-45°C to 50°C), and the accuracy of the sensors at these temperatures is poorly documented. In addition, the temperature corrections do not take into account the full temperature profile between the sensor and the surface. Under stable boundary layer conditions, the temperature gradient can exceed 25 K between the surface and the first 10 m (Vignon et al. 2017), and it is then recommended to define the temperature profile using multiple thermometers. When an operator is present, manual measurements can provide data for sensor validation. This data can be used to propose a calibration curve for the temperature dependence of RULS. However, this correction curve requires documentation of the full range of observable temperatures and manual measurements are not possible in remote areas in winter. In case of low temperature or high lapse rate in the surface boundary layer, the data must be flagged.

5.3.2.2   Impact of saltation / flotation

Blowing snow (saltation / flotation) is very common in Antarctica and can induce large errors (one or several tenths of cm) in elevation estimates because airborne ice grains are strong enough scatterers to reflect ultrasound waves. Probability of occurrence of blowing snow at the coast reaches 60% of the time, but it may exceed 80% of the time in the coast to plateau transition zone (Amory, 2020). The thickness of the saltation or flotation layer depends on wind speed and on the surface snow characteristics (snow hardness, density, and the surface roughness length) (Amory et al., 2021). Snow saltation and suspension are largely a turbulence problem and depend primarily on the friction velocity (u^* in m s^{-1}), i.e., the shear stress exerted by the flow at the surface. Snow erosion and saltation begin, when u^* exceeds the threshold friction velocity (u_t^*). The first step is thus to estimate u^*.

u^* is controlled by the wind speed and a drag coefficient, which can be calculated from measurements of the wind profile. The drag coefficient and the aerodynamic surface roughness length are interchangeable quantities for characterizing the aerodynamic properties of a surface (e.g., Andreas, 2011; Amory et al. 2017). However, "aerodynamic" roughness length, z0, is difficult to assess in the field, as it changes over time, depending on the history of erosion/deposition. z0 changes also depend on wind direction as surface sastrugis erode and adapt themselves according to the wind (Vignon et al., 2017). Andreas (2011) proposed that this variable could be estimated with the physical roughness length (which is a metric based on the wave-number spectrum of the snow surface elevation in the along-wind direction (Andreas, 2011)). But this relationship has not been verified over Antarctica to our knowledge. Moreover, retrieving this information is only possible with accurate lidar or multiple manual information, which is impossible to obtain except in summer during field work.

Concerning u_t^*, the uncertainties are even larger. This variable depends on the resistive gravitational and cohesive forces and on temperature (Schmidt, 1980) and surface snow metamorphism history (Gallée et al. 2001). Indeed, snow hardness mainly depends on inter-particle cohesion, density, grain shape and size (e.g. Amoy et al., 2020), which are related with the date of the last precipitation, on temperature of the atmosphere during previous days due to its impact on snow sintering (e.g., Amory et al., 2017). In other words, saltation depends on snow surface past history over days, to months on the Plateau (Picard et al. 2019). In theory, modelling saltation or suspension occurrences is possible, but it requires complex modelling approaches, accounting for drifting snow processes. This calculation is thus impossible in the field and should be done only ad-hoc.

The occurrence of a saltation layer is thus generally measured using sonic sensors for the aeolian Transport of Snow (as for instance Flowcapts©) (Trouvilliez et al. 2015). However, this sensor does not inform on the thickness of the saltation layer, and it is only accurate when not buried, justifying to install it a few centimetres or decimetres above the surface: in these conditions pure saltation events are missed.

Since saltation layer thickness is very complex to assess in the field, measurements of ultrasonic rangers are impossible to correct, and spurious data associated with saltation events are generally removed.

5.3.2.3   Range calculation algorithm

The manufacturers of ultrasonic gauges offer algorithms to calculate RULS from tfirst_echo. In the case of the SR50A gauge, this algorithm also provides echo quality information that can help eliminate spurious data in many applications. The quality numbers indicate the measurement uncertainty which may result from:

1) a sensor not perpendicular to the reflector;
2) a small or a poor reflector of sound; and
3) a rough target surface.

The measurement quality decreases with the occurrence of precipitation or blowing snow due to the presence of hydrometeors. Rain or mixed precipitation events could also impact the surface conditions and reduce the measurements certainty (SPICE final report, 2018).

Data processing by the algorithm is initially designed to identify "Reduced Echo" as "Highly Uncertain". However, it is unclear how to make use "Reduced Echo" criterium for the purpose of a quality control. During the WMO SPICE project, an independent data quality control process was proposed to evaluate "suspicious/not suspicious" data. This quality control was compared the "not good/good" data proposed by the "Reduced Echo" criterium. It was shown that the percentage of data described as "Good" by the sensor and then identified as "Suspicious" by the SPICE data quality control was very low (less than 0.3%) suggesting that the "Reduced Echo" criterium correctly excludes the suspicious data. However, data described as "Not Good" with the "Reduced Echo" criterium were frequently not flagged as "Suspicious" by SPICE data quality control. As a consequence, the criterion given by the manufacturer to remove data impacted by drifting snow events is insufficient in Antarctica (SPICE final report, 2018).

The Centre d'Etude de le Neige (CEN) of Météo-France and the IGE have proposed additional calculations to eliminate erroneous measurements (not published). This is studied through the realization of multiple chains of soundings (within a 10 s interval), allowing to flag data when the standard deviation σ_tfirst_echo exceeds a threshold. Then, a final analysis of the echo quality level allows a final selection of accurate data, which are not affected by saltation events (SPICE final report, 2018). If data are not removed, flagged data should be presented with a large uncertainty range to warn users on the high risk of using data for scientific purposes.

5.3.2.4   Range uncertainty

To sum up, ultrasonic gauges only inform on the first echo (tfirst_echo) due to surface features located below the sensor and multiple sensors are necessary to assess the average surface elevation over a few square metre areas. tfirst_echo variable is a function of air temperature (Tair), and on the presence (or not) of airborne snow particles between the sensor and the surface, i.e. if the thickness of the saltation or flotation layer (hsaltation/flotation) is not zero. After corrections and elimination of blowing snow artifact, the accuracy of one punctual RULS data is ±2cm. Optical sensors are much more independent of temperature, but are still (even more) impacted by snow particles between the surface and the sensor. Finally, small laser based measurement technologies, as the Multipoint Scanning Snowfall Sensor, could inform on more than one point below the sensor. This clearly does not represent the mean elevation of the surface in case of large surface roughness (sastrugis or dunes) and the impact of the surface slope at the scale of S3 footprint. The approaches to document elevation changes at larger scale are described hereafter.

5.3.2.5   Range variability in space and time

In Antarctica, the snow surface is perpetually changing even when no precipitation is occurring (e.g., Amory et al., 2016) due to wind that leads to snow erosion and non-uniform snow deposition. These non-uniform depositions take the form of over-accumulation "patches" having size in metres, more rarely in tens of metres. Most of these patches (60 %–85 %) are ephemeral but can survive a few weeks to months before being eroded (Picard et al., 2019) unless they stay definitely and are incorporated into the snowpack. As a consequence, elevation changes measured at one point may not reflect those at larger scales, in particular at the decametre/kilometre scales, which is of interest in term of S3

measurements (Libois et al., 2014). Figure a presents how snow surface may evolve locally in a few hours, without presenting changes at the larger scales. The presented surface features are known as sastrugi, barcanes, dunes, etc. (e.g., Filhol and Sturm 2015). Sastrugi and dunes are on continuous reshaping (Gallée et al. 2013; Libois et al. 2014; Amory et al. 2021).

At Dome C, changes are very frequent (every 3 days on average). In Figure b we observe that the measured daily changes in surface height (at Dome C) generally present a normal distribution, but a significant number of measurements present changes exceeding $\pm 2\,\sigma$ or $\pm 3\,\sigma$. These values generally correspond to sastrugis crossing the sensor footprint. RULS should consider these spatial differences which are related to the roughness length of the surface irregularities at the surface.
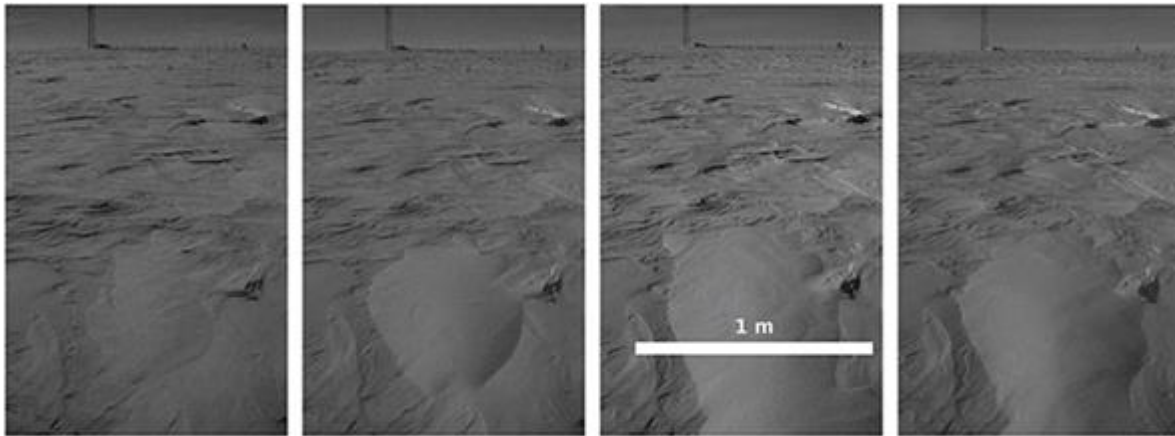
Figure 5.5 - A series of four pictures of the snow surface taken at a 1 h interval close to Concordia station on 29 November 2013, showing the formation of a thick snow patch whose thickness and horizontal extent are approximately 10 cm and 1 m². From Libois et al. (2014)
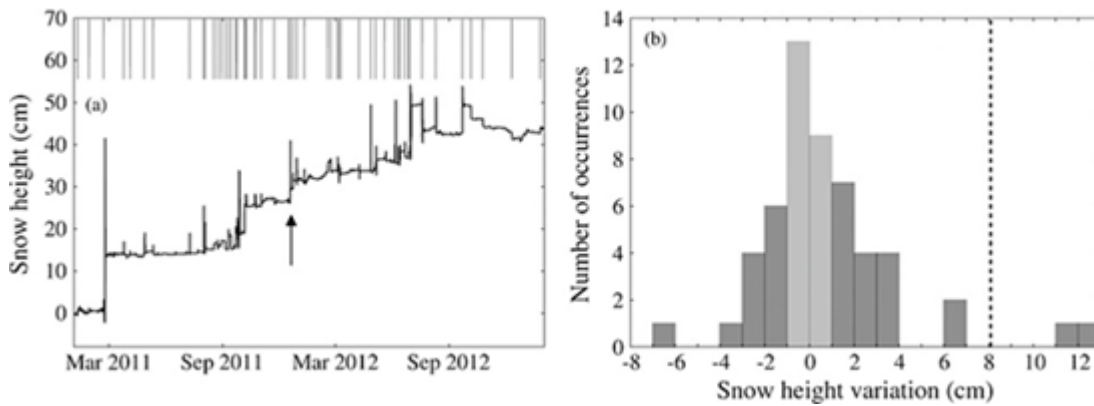
Figure 5.6 - (a) Hourly time series of snow height obtained from pictures of the snow surface. Vertical bars at the top highlight when drift events were detected from the snow height variations. The vertical arrow points to the drift event illustrated in Figure . (b) Distribution of snow height variations during the 53 drift events identified from pictures of the snow surface. The vertical dashed line indicates the mean annual accumulation from GLACIOCLIM stake network. From Libois et al. (2014).

As described in Smith (2014) review on the use of roughness in Earth sciences, this variable is not uniquely defined and could represent different properties of a surface (surface roughness), of a flow (flow resistance or roughness height) or of a model (calibration roughness). Amplitude-based definitions and metrics are quite common, and roughness parameters are generally associated with statistics of elevation distributions, including the median elevation above a reference level, percentiles of distribution, or the standard deviation of elevation. The latter represents a summary of topographic irregularity in units of length (Smith, 2017), but other statistics as the dominant frequencies of Fourier power spectra could be used to define the surface roughness, whereas wavelet transforms can also offer added benefit on space-frequency localisation of irregularities (Lampkin and Vanderberg, 2011).

Hereafter we describe different approaches to account for variations in the surface roughness, and in spatial differences in accumulation.

5.3.3 Relative surface height – optical rangers

To reduce the error due to the spatial variability of elevation to negligible levels, a scanning with a rough laser scan would be a great added value (Picard et al. 2016). Ground-based laser scanning devices for snowpack measurement have been developed and applied in the field (e.g. Prokop et al., 2008, Deems et al., 2013), in particular for continuous measurements (Adams et al., 2013; Picard et al., 2016, (Picard et al. 2019)). A rugged terrestrial laser scanning (RLS) device was proposed by Picard et al. (2016, 2019) for application at Dome C. This RLS was set up at 5.5 m above the surface and allowed scanning approximately approx. 200,000 points in 4 hours over 100-200 m2. This kind of device allows reaching an elevation estimation with an accuracy of 1 cm for a surface characterized by sastrugis of about 7-10 cm (RMSE) in height (Picard et al., 2016). This is to be compared to 10 cm with 10 ultrasonic sensors as estimated in the previous Section. With optimized power consumption, this sensor could be made fully autonomous (currently it requires 220V except one version operating in Canada with solar power + wind turbine). It could offer quantitative estimates of snow depth variations in remote areas with equivalent footprint and accuracy larger than a network of point gauges.

However, getting robust long-term measurements is difficult in remote areas where energy supply is complex to generate. Scientific SDMS-40 Multipoint Scanning Snowfall Sensor (referred to as MSSS hereafter) could provide 36 points of scans in a target area with a diameter of 30 cm – 200 cm. Nevertheless, its suitability and accuracy below -40°C has not been demonstrated yet. The MSSS would present large advantages in the field, in particular because the measurement is not impacted by temperature changes, and it allows getting height changes at numerous locations, Nevertheless, MSSS has never been used in tough environment as in Antarctica.

5.3.4 Relative surface height – sensor networks

A network of multiple sensors is necessary to accurately document surface changes. The installation of multiple snow height gauges (sonic rangers) in a 5x5 m2 area has been proposed in several occasions (e.g., Eisen et al., 2008, Picard et al. 2016, 2019) to make statistical analyses on RULS (Correlation, σ…) and offer information on data accuracy. To get information on the number of sensor necessary to measure the spatial variability of surface elevation over Antarctic snow, Picard et al. (2016) used data from a rough laser scan installed at Dome C, East Antarctica. For a surface RMS height of typically 7-10 cm, they demonstrated that a number N = 5 of punctual observations (with ultrasonic gauges or accumulation stakes) would be necessary to assess the elevation with an error of ≈ 3.5 cm (~40-50% of the RMSE). With N =10, the error would be 2.5 cm (25-35% of the RMSE). To attain an error of 1.5 cm (15-20% of the RMSE), 30 sensors would be necessary. For a different surface roughness, the relative uncertainty would be independent of the number of sensors, thus the number of sensors to get a similar absolute accuracy would significantly increase with higher surface roughness accordingly.

As a conclusion, except in extremely rough regions where sastrugis could exceed 50 cm, a network of 10 ultrasonic gauges would allow reaching an accuracy of about 10 cm or better.

5.3.5 Auxiliary observations – snow properties

The absolute surface elevation is the main target variable as a FRM for altimetric radar. It means that we consider here that the altimeter algorithm is aimed at providing the surface elevation, and the in-situ measurements described up to now are aimed at this purpose.

However, in the case of dry snow, which covers most of the Antarctic (>90% in summer and more in any other season), the penetration of the radar signal in the snowpack is a significant and well-known problem to estimate the surface elevation (Legrésy and Rémy, 1997). In order to get insight in the future (probable) discrepancies between the satellite data and the FRM, ancillary data can be collected. The penetration depth is mainly sensitivity to the absorption and scattering of the microwave in the snow. The temperature and density control the absorption whilst the grain size and density control scattering (Adodo et al. 2018). For this reason, we consider here the measurements of these three quantities.

5.3.5.1 Snow temperature profile

Several sites are equipped with automatic temperature probe strings (e.g. Dome C, Cap Prudhomme) up to depth of a few metres, which is sufficient for the altimetric penetration at C band and higher frequencies. The accuracy is largely

sufficient (typically 0.1°C) considering that the ice absorption coefficient is poorly constrained (Mätlzer et al. 2006) and especially its temperature dependency. As an alternative to the in-situ measurements, snow temperature can also be obtained using snow energy balance and thermal diffusion models (e.g. Picard et al. 2009) fed by reanalysis meteorological data which enable seamless simulations for the whole continent at scale of ~25km. The accuracy is of the order of 1-2 K (Fréville et al. 2014) which is sufficient with respect to the uncertainties on the ice absorption coefficient.

5.3.5.2   Snow density

Manual measurements of this quantity are usually considered as easy, the accuracy is typically better than 10%. In contrast, automatic measurements are very difficult since no direct method exists and the existing methods (e.g. dielectric measurements) rarely achieve 10% accuracy in real settings. As a consequence, many datasets provide surface densities for the summer near the stations and on the traverses (e.g. SUMUP database) but the temporal coverage is extremely poor and is biased towards the summer. As part of the GLACIOCLIM observatory, 1-m average density is measured in routine every summer to complement stack emergence height for the calculation of the annual snow accumulation rate (in kg/m2/yr). These data could be useful for the FRM application. Year-round surface density measurements are also taken at Concordia station (since 2011) with a twice-weekly repetitivity but this dataset is unique to our knowledge, and difficult to extrapolate as the Dome C is a peculiar windless slope-less place. Density is a highly dynamical variable and also is variable in space. Statistical modelling of its variations is difficult. Physical modelling is also difficult because of the number of poorly constrained involved processes (e.g. blowing snow as described above). Modelling is not a realistic alternative to in-situ measurements yet for FRM applications (Nanders et al. Submitted).

5.3.5.3   Grain size and and related metrics

The effect of snow microstructure on wave scattering is a complex matter, and grain sizes measured in the field, or more recently the Surface Specific Area (SSA, Arnaud et al. 2011) are commonly the metrics of choice to characterize the snow (Royer et al. 2017). Measurements of surface grain size can be done with automatic albedometer (Picard et al 2016) using the strong relationship between optical diameter and reflectance in the near and SWIR infrared. However, this technique is inadequate to measure the full profile of grain size as needed for radar penetration up to a few metres depth. Manual measurement is the only solution and has been done in Antarctica either samples by samples in snow pits or along snow cores (e.g. Picard et al. 2014) or using a specific profiling device (Arnaud et al. 2011). These measurements require more expertise and time than for density for instance, but constitutes a possible and realistic solution.

Regarding modelling, grain size is as variable as density and is difficult to predict as well, especially when considering that scattering is approximately proportional to the cubic power of the grain size (Mätlzer et al. 1998), putting a stringent requirement on the grain size uncertainties. High frequencies radars are the most sensitive to this variable.

6   Use of other satellite missions

This topic is covered by WP 3.2.3 about ICESat-2 and WP3.2.4 about high resolution DEMs, which are provided as supplemental technical reports to the roadmap TD-6. The reports include a review of ICESat-2 and high-resolution DEMs, and detailed analyses of their suitability for Sentinel-3 validation and improvements in Sentinel-3 processing.

7 FRM Protocols and Procedures

7.1 Metrological approach to uncertainties

FRMs, as described in section 1.3, are non-satellite observations that meet the standards of QA4EO and which are helpful to the calibration and validation of satellite observations. For FRMs to meet the standards of QA4EO, they need to have traceability to a community-agreed reference (preferably SI) and have an associated quality metric (preferably a robust uncertainty budget). Several different projects have applied these principles to a wide range of satellite and non-satellite observations, and from those projects, guidelines have been established that are documented on the QA4EO website (www.qa4eo.org).

These guidelines set out 5 steps to an uncertainty budget, which are discussed in the subsections below.

7.1.1 Step 1: Define the measurand and the measurement model

Defining exactly what is being measured and provided in a data set is often more difficult than it first appears. Even for an in-situ observation, the reading on the instrument (e.g., a temperature) may be different depending on how the measurement is made, i.e., on its input quantities (e.g., through a radiance measurement, or the expansion of mercury or the resistance of a thermocouple). And beyond that, the measurand of interest may be in how that reading relates to the estimate of the underlying physical phenomenon (e.g., air temperature near surface), or some representative phenomena (e.g., average air temperature in a grid cell of a model). Similarly, for a satellite observation, the measured signal, often in 'counts' needs to be converted to a physical quantity (e.g., top-of-atmosphere radiance within a spectral band). Processes such as orthorectification alter the perception of the measurand. Is an observation representing an average value within a pixel, or a peak value within a footprint? When satellite and in situ data are compared, they are likely to measure different things. For example, satellite-based measurements may relate to sea surface temperature as the top micron of the water, measured over a satellite footprint, whereas in-situ measurements may relate to sea surface temperature at a single point at a depth of a few tens of centimetres.

Furthermore, there may be questions of reference – is a range measured relative to the Earth's ellipsoid or to its geoid, for example. At higher levels of processing, where measured values are combined with models, the measurand may be even more difficult to define. However, defining the measurand is important both to describe the dataset to users and to enable clear thinking in the uncertainty budget. Sometimes, it is necessary to do separate uncertainty analysis for different linked measurands and propagate uncertainty between these steps (e.g., the uncertainty associated with the point temperature, the uncertainty associated with the average spatial cell temperature assumed from that point, and the uncertainty associated with comparing the in-situ temperature to the satellite temperature all require separate analysis).

The measurement model itself may be able to be written as an equation with an analytical function. Or it may only be defined through code, particularly if iterative processes, non-linear fitting, or machine learning techniques are part of the processing. Whether or not it can be written as an equation, the processes by which input quantities are combined to determine the measurand, is known as the measurement model. It is important to realise that there will be uncertainties associated with the form of the measurement model (whether the process it describes accurately describes reality) as well as with the input quantities that are used within it.

7.1.2 Step 2: Establish traceability with a diagram

A visual representation of how a measurement and its traceability is achieved, along with visually representing the different sources of uncertainty, are highly valuable in assessing performance. Diagrams are extremely useful mind mapping tools to help understand and communicate how a measurand is derived and to consider and share what the sources of uncertainty are. Diagrams show where terms come from and thus highlight sources of uncertainty in input quantities and in the approximations and assumptions inherent in the model.

There are different types of diagrams that can be helpful for different purposes. In this project we use a combination of uncertainty tree diagrams and more traditional processing diagrams, as well as a new concept of a comparison diagram.

7.1.3   Step 3: Evaluate each source of uncertainty and document in an effects table

After the work in step 1 to specify the measurand, and in step 2 to identify where the input quantities of the measurement model all come from, it should be possible to get a list of sources of uncertainty (also known as effects). There are several things that need to be known about each effect and FIDUCEO and GAIA-CLIM used the concept of an 'effects table' to document, systematically, the information that needs to be known about each effect.

The exact rows of an effects table will depend on the application, but there are several common requirements. This is for each source of uncertainty to identify:

- Which quantity in the measurement model it affects?
- The magnitude of the uncertainty
- The shape of the probability distribution function for the uncertainty
- How the uncertainty associated with this effect is propagated to the measurand (the sensitivity coefficient)

The error correlation shape and scale for all 'dimensions' that are relevant both for determining the measurand and for subsequent 'higher level' processing or applications that perform averages and/or comparisons.

Additionally, it is valuable to document whether the analysis in the table is mature (based on sound analysis with evidence and validated through independent comparison) or very immature (based on expert judgement) or somewhere in between.

The 'effects table' provides a common method for recording what is known about each source of uncertainty. This is valuable to think through the uncertainty analysis and for recording for long term data preservation purposes. Using effects tables that follow the documentary templates and examples given in the guidelines will lead to consistency within the community.

A core part of this methodology (central to step 3) is to consider error-correlation shape and scale in all dimensions. This requires a careful distinction between the concepts of 'uncertainty' and 'error' and an understanding of the nature of environmental observations (affected by both instrument uncertainties and natural variability) and the nature of satellite observation data processing (with different 'levels' that are often processed by different scientific communities). A full discussion of these aspects is given in the documents on the www.qa4eo.org website.

7.1.4   Step 4: Calculate the FRM and its associated uncertainty

The next step involves processing the FRM, FDR or TDP through the measurement model and determining the associated uncertainties. These uncertainties need to be propagated all the way through the entire processing chain to the measurand.

There are two ways of processing uncertainties that are described in the GUM (Guide to the expression of uncertainty in measurement, BIPM, 2008). Uncertainties may be processed using Monte Carlo methods (as in an ensemble analysis), or through the Law of Propagation of Uncertainties (a linearized Taylor expansion often recognized as 'the square root of the sum of the squares', although when there is error correlation a full covariance matrix is needed). Monte Carlo can provide better results for non-linear models and is the often the only option where the processing cannot be written analytically (e.g., in neural networks or iterative processes), however, it is computationally expensive and does not provide easy access to the importance of different sources of uncertainty. A hybrid approach can use Monte Carlo analysis to evaluate sensitivity coefficients that are propagated through the law of propagation of uncertainties or used in look up tables.

7.1.5   Step 5: Document for different purposes

Record the information for both today's users (simplified summary information that can be readily used by others) and for the purposes of long-term data preservation (recording and documenting all the information needed for the scientific analysis to be reproducible in the future).

7.1.6   Work so far in this project

In this project our most significant work has been in the first two steps of the approach to uncertainty analysis. Considerable thought has gone into the careful definition of the measurand, documenting in equations exactly what is measured, and how that differs from location to location. Uncertainty traceability diagrams have been drawn for the

fixed station measurements and the satellite measurements, to show the traceability (step 2). Some indication has been given about how uncertainties can be quantified (step 3), though that work must be expanded in phase 2 of the project.

7.2 Overview of the metrological analysis and comparison process

7.2.1 Comparison aims

An overview of the comparison process is given in Figure 7.1. The FRM that is compared with the satellite is from an average of multiple instruments as part of the fixed station (#1 – #7 on right), which is then averaged over an hour and then position and time matched to the satellite overpass. Details are given in the sections below.
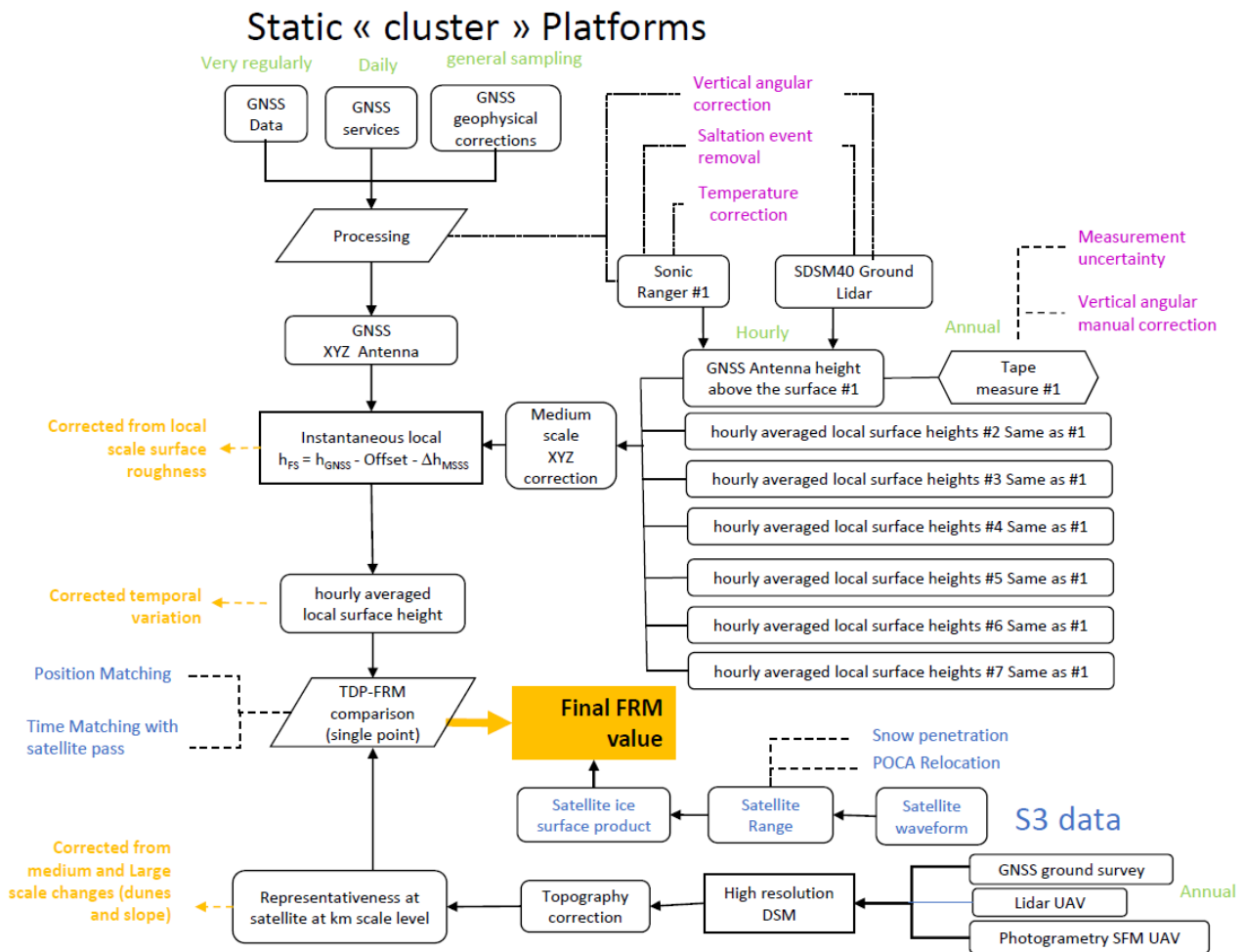
Figure 7.1: Schematic of a static platform for FRM

7.2.2 The surface that is measured

As discussed in Section 3.1, land-ice has an evolving surface that consists of layers of ice, multi-year snow (firn) and new snow on top of more stable bedrock. For the purposes of uncertainty analysis, we simplify this to consider three basic layers: an underlying bedrock terrain that changes only very slowly (>1 year), covered by an intermediate ice and firn level, that changes over days to months, and a top snow layer that erodes and deposits by the wind (and by melting and precipitation) and changes in minutes and hours.

Annex A, provided as Section 8, discusses how to think about the natural variability of this real surface. Ideally, to compare with the satellite, we would have ground stations at the points of the satellite POCA, at the time of the satellite overpass. This is not possible. Instead, we combine a fixed station measurement, which gives an absolute height reference at its own location and can provide a time series that describes intermediate level changes, with a DSM that enables the fixed-station measurements to be transferred to the POCA of an individual overpass.

As discussed in Section 8, these measurements, even in the absence of measurement uncertainty from the instruments (which will also be there), will fluctuate due to the top wind-swept layer that changes in minutes and hours, and has a spatial scale of decametres. To limit the impact of these fluctuations, the fixed station measurements are averaged in both time and space – with several measurements a few metres apart to sample the decametre scale variation, and measurements over an hour or so to sample the temporal variability.

A set of individual comparisons between a single satellite observation and the fixed station, even with the fixed station averaging, will also fluctuate from satellite observation to satellite observation, due to the natural variability from one POCA point to another (temporally and spatially). For this reason, it's important to average over multiple POCA observations.

In the following sections we present the uncertainty analysis for the satellite measurements (Section 7.3), for the fixed station measurements (Section 7.4) and for the transfer of the fixed station measurement to the POCA for the satellite and the comparison process (Section 7.5). We only briefly touch on validation campaigns through airborne or ground-based surveys here (Section 7.4.3) as that is described in more detail in the related Sea Ice report [TD-2] with many of the same principles and uncertainties applying to land ice, as well.

7.3   Metrological uncertainty analysis for Sentinel-3 data products

To ensure a full metrological uncertainty analysis, the uncertainty traceability must be performed for both the satellite observations and for the FRMs when a comparison is done. Here, we present an example of a metrological uncertainty analysis for the most relevant data product for land ice applications, the level-2 land-ice thematic product. We briefly describe the processing chain from initial satellite observations to post-processed data product for surface elevation, the measurand to be compared with FRMs, as well as other processes that impacts the uncertainty of the data product.

7.3.1   The Sentinel-3 Land Ice Thematic Product

In addition to the operational Sentinel-3 STM LAND level-2 products, thematic data products (TDP) are generated for areas of (1) inland waters, (2) sea ice, and (3) land ice [RD6]. This example concerns the land ice thematic product, known as *SR-2-LAN-LI*, but also applies more generally to the standard LAND product. The main improvements for the land ice product are an extended window processing to sample more of the SAR waveforms over rough/sloping terrain (Aublanc et al 2018) and a finer slope-model for Antarctica and Greenland (Helm et al. 2014) to relocate the measurements more precisely to the point of closest approach (POCA) to the satellite.

7.3.1.1   Defining the measurand and measurement model

Radar altimeters are active sensors emitting microwave pulses and measuring the backscattered energy as a function of time, providing the so-called waveform. Satellite-to-surface range is deduced from this two-way travel time, once the surface echo has been precisely identified in the waveform, the process known as surface retracking (e.g. Rémy and Parouty, 2009). Surface ellipsoidal height (h) is then derived by subtracting the measured range (ρ) from the satellite height (h_{sat}), while correcting for atmospheric signal delays (C_{atm}) and tidal effects on land (C_{tides}). If the surface within the SAR footprint is not flat, the measurement point should also be relocated from satellite nadir (φ_0, λ_0) to the satellite point of closest approach (φ, λ), in which case a topographic height correction (Δh_{topo}) is needed to account for the off-nadir slanting of the measured range (see Section 2). The measurement model can be summarized as:

$$h(\varphi, \lambda, t) = h_{\text{sat}}(\varphi_0, \lambda_0, t) - \rho(\varphi, \lambda, t) - \Delta h_{\text{topo}}(\varphi, \lambda) - C_{\text{atm}}(\varphi_0, \lambda_0, t) - C_{\text{tides}}(\varphi_0, \lambda_0, t) + 0. \quad (7.1)$$

where t represents the time of a satellite overpass at latitude φ_0 and longitude λ_0. The geodetic coordinates and heights are referenced to the WGS-84 ellipsoid. Satellite observations are sampled at a rate of 20 Hz, whereas correction terms are posted at 1 Hz and need to be resampled to the 20 Hz frequency of observations. The observational spacing along the ground track then becomes about 330 m, corresponding to the along-track SAR mode footprint resolution.

Satellite location and altitude are derived from a Precise Orbit Determination (POD) system which combines positioning measurements from two GNSS receivers using the higher altitude GPS and GALILEO satellites, a Doppler orbit radio positioning instrument (DORIS) that receives signals from DORIS beacons on the ground, and a Laser Retro Reflector (LRR) that facilitates precise laser ranging of the satellite from the ground. The POD orbital accuracy is stated to 3 cm (1-sigma RMS) for Non Time Critical (NTC) products of latency 28 days [RD7], like the land-ice thematic product.

An onboard surface tracker, operated in open loop or closed loop (Section 2), is used to estimate an approximate surface range (ρ_{trk}) to which the signal receiving window is centred at sample-gate 43 out of 128 in total, after extended window processing with 512 samples for the land-ice product [RD8]. The surface is then re-tracked more precisely in the waveform (Fig. 2.3) by an empirical algorithm or a physical model that is matched to the waveform. The two retrackers employed in the land-ice product are OCOG-/ICE-1 (Bamber et al., 1994) and UCL-ice-sheet (Wingham and Wallis, 2010), and ranges/elevations are included for both. Considering a derived time delay with respect to the initial tracker, the precise satellite-surface range can be described as:

$$\rho(\varphi, \lambda, t) = \rho_{\text{trk}}(\varphi, \lambda, t) + \frac{c\tau}{2} + C_{\text{instr}}(\varphi, \lambda, t) + 0, \qquad (7.2)$$

where C_{instr} is a net instrumental correction (C_{instr}) for ultrastable oscillator (USO) drift, internal path delay, distance to antenna centre-of-gravity (COG), modelled instrumental errors, system bias, as well as Doppler nadir and slant-range corrections [RD7]. The re-tracked range refers to the satellite point of closest approach (POCA, Figure 2.4) which means that the measured range is slanted with respect to nadir, causing a geometrical effect that needs to be accounted for in the altimetry equation (Eq. (7.1)) with a topographic height correction (Δh_{topo}). For the land-ice product, this correction is estimated from a linear surface-slope gradient (s) derived from an external DSM (Helm et al. 2014) considering a footprint-scale area that ideally extends form nadir (φ_0, λ_0) to POCA (φ, λ):

$$\Delta h_{\text{topo}}(\varphi, \lambda) = \frac{s^2 \rho}{2\left(1 + \frac{\rho}{R}\right)} + 0, \qquad (7.3)$$

where R is the Earth radius. The DSM and associated linear-slope model are used to relocate the altimetry measurement point from nadir (φ_0, λ_0) to POCA (ϕ, λ), in effect shifting the location across-track in the up-slope direction within the SAR footprint. The along-track slope has relatively little impact as the SAR footprint is much narrower along-track (~330 m) than across-track (several kilometres). The POCA offset correction can estimated from:

$$\begin{pmatrix} \phi \\ \lambda \end{pmatrix} = \begin{pmatrix} \phi_0 \\ \lambda_0 \end{pmatrix} + \frac{s\,\rho}{1 + \frac{\rho}{R}} \begin{pmatrix} \alpha_{\text{lat}} \\ \alpha_{\text{lon}} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad (7.4)$$

where the slope gradient (s) needs to consider the non-equidistant nature of geodetic coordinates – via α_{lat} and α_{lon} – or be used in a metric projection before converting to geodetic coordinates in the end. Note that POCA corrections are only carried out for the Antarctic and Greenland ice sheet (as per March 2023) where the CryoSat-2 based DSM of Helm et al. (2004) is available. In all other glacier regions, surface elevations are referenced to the nadir location and will therefore have large topography-dependent errors as the actual re-tracking will still occur at the unresolved POCA location.

Geophysical height corrections are applied by default to the land-ice product. The atmospheric correction (C_{atm}) in Eq. (7.1) consists of a wet troposphere correction, a dry troposphere correction, and an ionospheric correction:

$$C_{\text{atm}}(\varphi, \lambda, t) = C_{\text{wet trp}}(\varphi_0, \lambda_0, t) + C_{\text{dry trp}}(\varphi_0, \lambda_0, t) + C_{\text{iono}}(\varphi_0, \lambda_0, t) + 0. \qquad (7.5)$$

The ionospheric delay over land surfaces is estimated from the JPL Global Ionosphere Maps (GIM, Iijima et al. 1999), whereas the dry and wet tropospheric corrections are derived from vertical integration of ECMWF meteorological data.

The tidal height correction (C_{tides}) in Eq. (7.1) has three components over grounded land-ice:

$$C_{\text{tides}}(\varphi, \lambda, t) = C_{\text{Earth tide}}(\varphi_0, \lambda_0, t) + C_{\text{pole tide}}(\varphi_0, \lambda_0, t) + C_{\text{ocean loading}}(\varphi_0, \lambda_0, t) + 0. \qquad (7.6)$$

These corrections rely on tidal models, namely the Cartwright model for solid Earth tide (Cartwright and Edden 1973), the Wahr model for polar tide (Wahr 1985), and the FES-2014 (Lyard et al. 2006) model for ocean loading. Over ice shelves, the floating extensions of the Antarctic and Greenland ice sheets, additional corrections need to be applied for the ocean tide and the inverse barometer effect (IBE), as for the sea-ice thematic product. These two additional corrections are available in the land-ice data product but needs to be applied by the user when needed. The ocean tide correction stems from the same model as the ocean loading correction (FES-2014), whereas the IBE correction is derived from local atmospheric pressure using ECMWF meteorological data, with a 1 cm height correction for a change in pressure by 1 mbar (Padman et al. 2003). In the case of ice shelves, Eq. (7.1) can be expanded to:

$$h_{\text{iceshelves}}(\varphi, \lambda, t) = h(\varphi, \lambda, t) - C_{\text{Ocean tide}}(\varphi_0, \lambda_0, t) - C_{\text{IBE}}(\varphi_0, \lambda_0, t) + 0. \qquad (7.7)$$

Usually, the ellipsoidal heights (lowercase h) are used directly, but for some purposes it is more practical to have orthometric heights (uppercase H) that refer to a geoid that is close to mean sea level. The geoid model in Sentinel-3 data products is EGM2008. Orthometric height is derived by subtracting geoid height (N) from the ellipsoidal height:

$$H(\varphi, \lambda, t) = h(\varphi, \lambda, t) - N(\varphi, \lambda) + 0. \qquad (7.8)$$

The zero term in these measurement equations refer to remaining unquantified errors and uncertainties. The two main ones to consider over land ice is the POCA relocation and volume scattering (penetration) in snow.

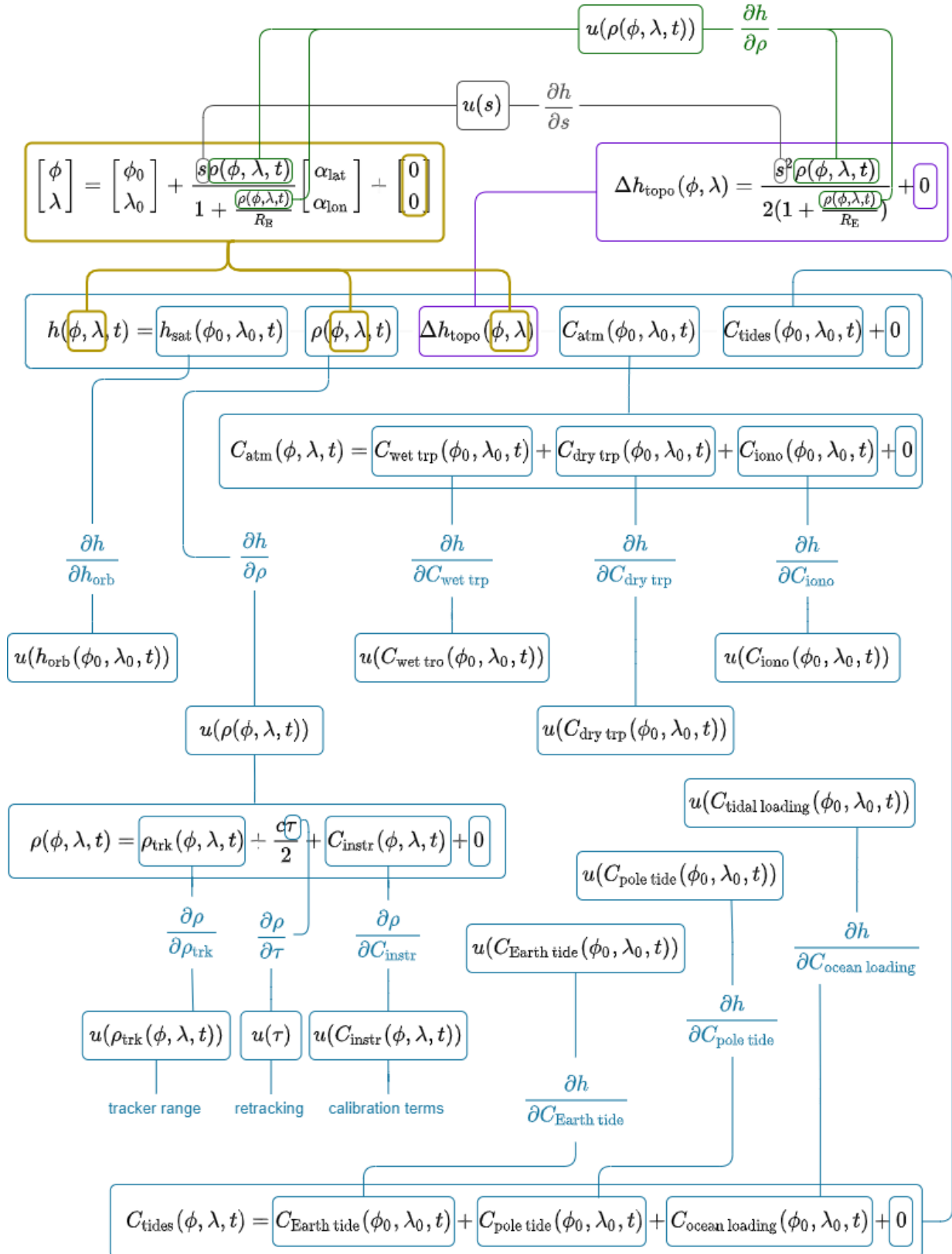### 7.3.1.3 Traceability uncertainty diagram

Figure 7.2 TDP traceability diagram for land ice

7.4 Implementing metrological uncertainty analysis for land ice FRMs

An FRM setup for land ice validation involves the application of multiple instruments. Therefore, in this section, we analyse the metrological details of various FRM components, either independently or in combination to one another. We aim at defining the measurands and measurement functions of these components and linking them to a traceability diagram. This shall cover the first two metrological steps to uncertainties (7.1).

For some components, we briefly touch upon the concept without going deep into details. This is done deliberately to reflect the complexity of an FRM setup for land ice validation. The current gaps in our analysis could be addressed in the follow-on of St3TART, if any.

7.4.1 On fixed stations

Multipoint Scanning Snowfall Sensors (MSSS) and sonic gauges (or rangers) are two of the common instruments to monitor height variation over land ice (see Section 5.3). The MSSS covers an area of the size of a few m^2. At every instance, the covered area is sampled at 36 positions. When using sonic gauges, it is necessary to employ 7 to 10 instruments within a confined area for a comparable coverage to MSSS. In this case, 7-10 measurements are made at every instance of time.

Sonic rangers are robust and have been tested over more than 25 years. They are currently the most appropriate sensor for FRM in Antarctica. The MSSS, however, has never been tested over long time periods. For sake of simplicity, we use the term MSSS in the rest of the document, but this name refers here to either a system of 7 sonic rangers or to 36 height measurements obtained with a MSSS. Here we take the example of an MSSS station accompanied by a GNSS receiver to collect in situ height measurements. However, with slight modifications, the following procedure should hold for the case of sonic gauges as well.

7.4.1.1 Defining the measurand and measurement model

As suggested by Figure 3.4, an MSSS acquires multiple sensor-to-surface range measurements with the specific sampling rate of τ'. Each individual sampling location may be noted as $(\phi', \lambda') \in \Omega'$. In Annex A (Section 8), the notation i is used for an individual measurement by an individual sensor. Here, Ω' describes the full area covered by the MSSS instrument or the set of sonic rangers. The GNSS receiver is also located within or very close to Ω'. The recorded range measurements are averaged over the spatial extent of Ω' and the temporal extent of $\delta\tau$ to filter out both spatial and temporal noise. As discussed in Annex A (Section 8), this averaging reduces the effect of the rapidly varying wind-blown surface. It is important to notice that the integration time is chosen considering several parameters. Here we assume that the integration time is selected as such that the final time series of range measurements at the location of MSSS can be represented by τ (eq. (7.9)).

$$\overline{\Delta h}_{\text{MSSS}}(\tau) = \underset{(\phi', \lambda') \in \Omega', \tau' \in \left[\tau - \frac{\delta\tau}{2}, \tau + \frac{\delta\tau}{2}\right]}{\text{mean}} \Delta h_{\text{MSSS}}(\phi', \lambda', \tau') \qquad (7.9)$$

In this setup, the absolute position of the fixed station is recorded with a geodetic GNSS receiver using PPP technique. It is important to notice that the GNSS system works at a different sampling rate, η'. To get more accurate GNSS positions, the measurements should be averaged over several hours, $\Delta\eta$:

$$\bar{h}_{\text{GNSS}}^{\text{PPP}}(\eta) = \underset{\eta' \in \left[\eta - \frac{\Delta\eta}{2}, \eta + \frac{\Delta\eta}{2}\right]}{\text{mean}} h_{\text{GNSS}}^{\text{PPP}}(\eta') \qquad (7.10)$$

According to (eq. (7.10)), $\bar{h}_{\text{GNSS}}^{\text{PPP}}(\eta)$ is the ellipsoidal height of the GNSS antenna at time η. As MSSS and GNSS do not share the same sampling properties, it is necessary to resample both measurement series to the new time vector $\boldsymbol{t}$. The sampling interval for this new time vector can be similar to that of the averaged MSSS measurements $\boldsymbol{\tau}$, or that of the satellite altimeter data products, e.g., 20 Hz. As suggested by Figure 7.3, the resampling will affect the uncertainty of measurements, and this shall be considered.

If the distance between GNSS antenna and MSSS, Δh_{offset}, can be obtained using a tape measure, the height of the land ice area covered by the fixed station can be calculated as:

$$h_{\text{FS}}(t) = \bar{h}_{\text{GNSS}}^{\text{PPP}}(t) - \Delta h_{\text{offset}} - \overline{\Delta h}_{\text{MSSS}}(t) + 0 \qquad (7.11)$$

Here we have not considered the uncertainties associated with the fixed station instrument measurements themselves. As discussed in the Annex A (Section 8), many sources of uncertainty, especially those with a random distribution of

errors, will be indistinguishable from natural variability. However, more systematic, or partially correlated effects will be important. For sonic stations, one of the most important sources of uncertainty is due to the temperature sensitivity of the speed of sound in air. At some Antarctica stations, the temperature of the air can change considerably, even over the small number of metres between the sonic sensor and the ground. This effect needs further investigation in a later phase of this project.

7.4.1.2   Establish traceability diagram for the FRM
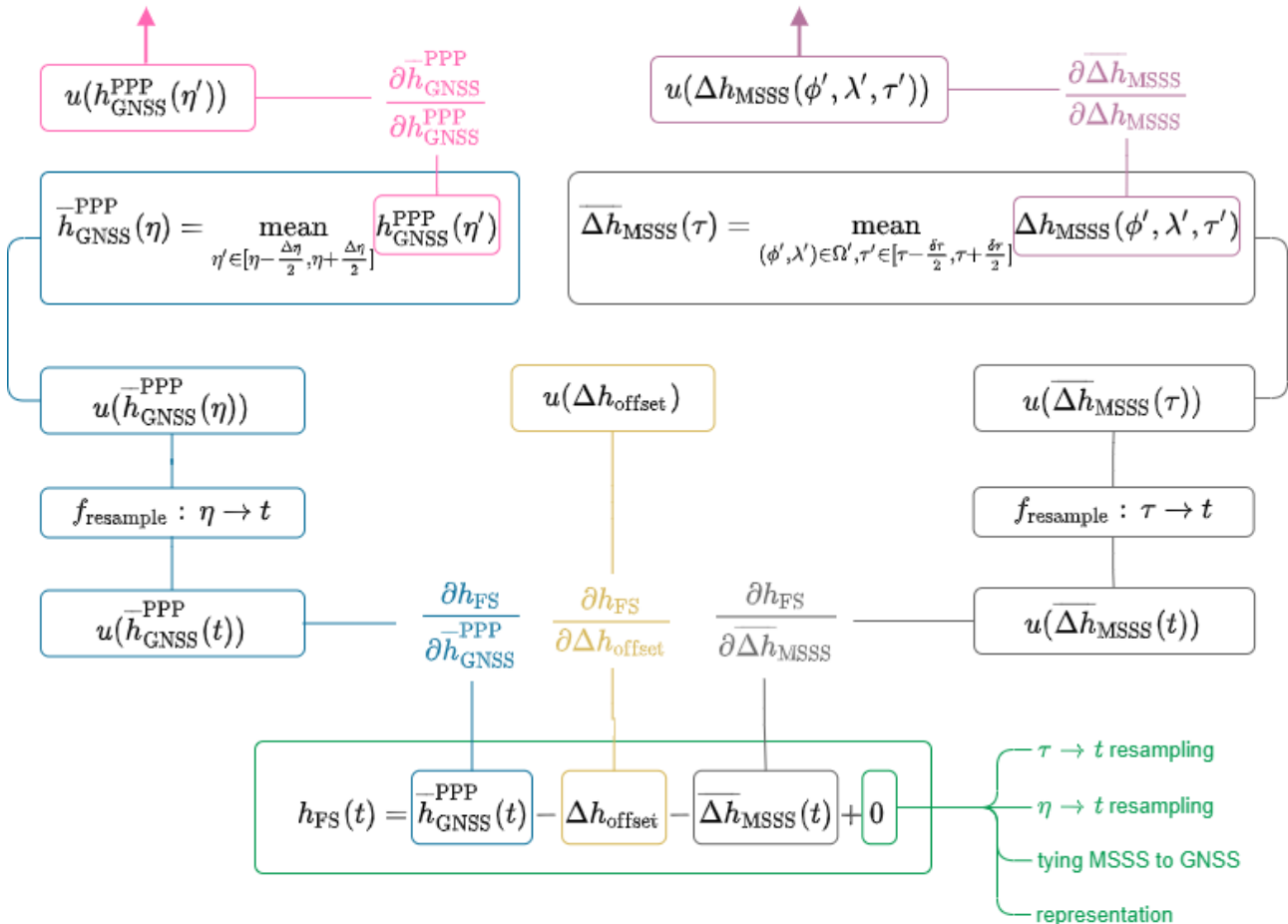
Figure 7.3 FRM traceability diagram for MSSS fixed station

7.4.2   On DSM from airborne photogrammetry

The FRM requires a DSM, in this example obtained from UAV photogrammetry, but could also be made from UAV lidar or a ground-based GNSS survey. A review of UAV photogrammetry is given in Section 5.2.2. For the purposes of the analysis, the relative DSM obtained from UAV photogrammetry provides the offset in height between the fixed station location $(\phi_{FS}, \lambda_{FS})$ and the location of the POCA for any one satellite observation, observation k for orbit j. The DSM needs to connect the fixed station to the POCA sampling locations, for example with a survey that covers a few kilometres to either side (Figure 4.1), This station-to-POCA height difference is valid for the time of the measurement, but as described in Annex A (Section 8), the assumption is made that any changes between that time and the time of the satellite overpass of interest is either due to the windswept layer that is rapidly changing (and therefore reduced by averaging), or due the intermediate layer and averaging out for multiple POCAs.

Uncertainties associated with the DSM measurements themselves will include three types of components that mirror the natural variability. Any noise is expected to average out over the surface. Any fully systematic effect (an absolute offset) will not be important because it will affect the POCA and the fixed station equally and be removed by the double difference. The most significant uncertainties will be those that create a slope across the surface or a relative height

change between the fixed station and POCA locations. Effects in the DSM will be difficult to distinguish from real slope effects in the changing landscape and can be analysed as discussed in Section 8.9.

An indicative value for the uncertainty associated with the DSM can be obtained from other studies that have compared UAV-borne photogrammetry systems with in situ measurements taken at the same time (see Section 5.2.2).

7.4.3 On absolute validation campaigns

As described in section 4.1, it is possible to make an FRM from either a relative DSM combined with a fixed station (as described above), or to use an absolute FRM from a larger scale ground-based or airborne campaign. In the related Sea Ice report [TD-2], the uncertainties associated with airborne lidar and/or a Ku-band radar instrument are described in detail. Many of the same uncertainties are relevant over land ice.

7.5 Comparison

Figure 3.4 provides an overview of the radar altimetry measurements, where the actual altimetry track at cycle j, is deviated from the theoretical track by a few hundred metres. The altimetry measurements are distributed with sampling rates of 1 Hz and 20 Hz. For our analysis, we take the 20 Hz measurements. The POCA samples do not follow the nadir track. Depending on the local topography within the SAR footprint, the POCA location for every single 20 Hz sample is going to differ from the nadir location. We denote the spatiotemporal properties of the POCA samples by $(\phi_{k,j}, \lambda_{k,j}, t_{k,j})$ and the corresponding nadir sample with $(\phi_{0_{k,j}}, \lambda_{0_{k,j}}, t_{k,j})$, where j defines the cycle number and k points to a specific discrete sample during cycle j. The corresponding POCA and nadir samples share the exact same time $t_{k,j}$.

Different methods for comparison between land ice FRMs and TDPs are conceivable. Figure 3.4 points to some of these methods by highlighting the possibility of defining a circular area around a fixed station or drawing attention to narrow bands that resemble SAR footprints. A comparison may even be designed over larger areas, by gridding the spatial dimension or by dedicated surveying along satellite ground tracks. Here, we focus on a comparison within a confined area and only briefly touch upon comparison over larger areas.

7.5.1 Comparison within a confined area

In this scenario, a fixed station is set up at some location using a GNSS receiver with either a Multipoint Scanning Snowfall Sensor (MSSS) or a set of sonic rangers. Any range measurement from the MSSS or the sonic rangers can be transferred to a height measurement using the local height reference from the GNSS. We denote the extended area of the fixed station as Ω' (Figure 3.4). Ω' contains the GNSS and the MSSS (or the sonic rangers). To compare the altimetry measurements of height to that of the fixed station, a DSM that covers both is needed. Given a DSM, it is possible to correct for the height difference between each 20 Hz POCA sample and the fixed station. As depicted in Figure 4.1, the local DSM can be derived from a campaign around the fixed station, where:

- o photogrammetry is used to derive a DSM,
- o a lidar instrument scans the area, and the acquired point cloud is used to derive a DSM, or
- o a GNSS receiver on a moving platform is used to survey the area and interpolate into a DSM.

Comparison within a confined area fundamentally resembles the concept of defining a virtual station which is common for inland altimetry applications. In this example, a fixed station is installed close to the theoretical orbit (or predicted POCA sampling) of Sentinel-3 and a local DSM of the area is generated. The concept is to validate all satellite measurements which fall inside a reference area around the fixed station, here taken as a circular area of a 2 km radius as shown in Figure 3.4. A larger radius may be required for more complex or sloping terrains (the POCA separation from nadir can be estimated from Eq. 7.4). The dashed lines in the figure show the nadir trajectory of the satellite during overpass j. However, due to the topography variation within the SAR footprint, the leading edge of the waveform return come from the point of closest approach, the POCA. The location of every POCA measurement is noted as $(\phi_{k,j}, \lambda_{k,j})$, where k is the sample index of the overpass j.

According to Figure 7.2, the height of land ice at each POCA location can be described as $h(\phi, \lambda, t)$. This quantity (see also section 8.2), is measured by the satellite altimeter to give $h_{\mathrm{ALT}}(\phi, \lambda, t)$. To distinguish from one POCA to another, as shown at the top section of Figure 7.4, we use a more accurate notation of $h_{\mathrm{ALT}}(\phi_{k,j}, \lambda_{k,j}, t_{k,j})$ to describe this measured value. On the other hand, according to Figure 7.3, the measured height collected at the fixed station at time t is described as $h_{\mathrm{FS}}(t)$. In this section, we use a more accurate notation of $h_{\mathrm{FS}}(\phi_{\mathrm{FS}}, \lambda_{\mathrm{FS}}, t_j)$ to refer to the same

measurement at the fixed station. Please note that the time sample of altimetry measurements has two subscripts of j and k to let us distinguish between samples which are taken at different locations along a single track. For the fixed station, however, the time variable has only one subscript, j. This is because we can assume that within a fraction of a second – when the satellite enters and exits the boundaries of $\Omega_{2\text{km}}$ – the measured height at the fixed station does not change. In effect, we are assuming that

$$t_j = \underset{\forall k \in j}{\text{mean}}\, t_{k,j}\,. \tag{7.12}$$

The other source of information is a local DSM model. For point x on the DSM, the height of DSM is $h'_{\text{DSM}}(\phi_{\text{x}}, \lambda_{\text{x}}, t_{\text{c}})$. Here, we use a slightly different notation for height, h', as there is no strict requirement that the DSM is referenced to the same ellipsoid as the TDP and the collected heights at the fixed station. We use the DSM to correct for the height difference between two points. The temporal characteristic of the DSM is also different. The DSM is derived from a specific campaign at time t_{c}, where c stands for 'campaign'.

Given the measurements at the fixed station, a local DSM, and the satellite measurements, we may compare TDP to FRM with two different approaches.

7.5.1.1   Approach 1

In the first approach, we assume that all measurements of the fixed station have been resampled to the same time stamps as the TDP, e.g., the 20 Hz time stamp. In this case, we can define an FRM measurement that can be compared to a single TDP measurement at the POCA. In fact, we need to take the height measurement of the fixed station at time of overflight j, and transfer it to the POCA location of $(\phi_{k,j}, \lambda_{k,j})$ using the DSM model:

$$h_{\text{FRM}}(\phi_{k,j}, \lambda_{k,j}, t_j) = h_{\text{FS}}(\phi_{\text{FS}}, \lambda_{\text{FS}}, t_j) - h'_{\text{DSM}}(\phi_{\text{FS}}, \lambda_{\text{FS}}, t_{\text{c}}) + h'_{\text{DSM}}(\phi_{k,j}, \lambda_{k,j}, t_{\text{c}}) + 0.. \tag{7.13}$$

Here, the FRM height is specific to each altimetry sample in terms of its spatial properties, $(\phi_{k,j}, \lambda_{k,j})$. However, the timing, t_j, for each sample k is not unique: the temporal transfer for the comparison with the fixed station happens at the resolution of t_j and not $t_{k,j}$. In the next step, we can directly compare the TDP to the generated FRM at each sampling point,

$$r(\phi_{k,j}, \lambda_{k,j}, t_{k,j}) = h_{\text{ALT}}(\phi_{k,j}, \lambda_{k,j}, t_{k,j}) - h_{\text{FRM}}(\phi_{k,j}, \lambda_{k,j}, t_j) + 0, \tag{7.14}$$

where $r(\phi_{k,j}, \lambda_{k,j}, t_{k,j})$ is difference between the satellite and FRM at a specific location and instance of time. Figure 7.4 combines the above equations in one.

To get a more reliable understanding of the difference of the Sentinel-3 height estimates to the FRM, we can then average all such residuals within the confined area and for as long as corresponding measurements of satellite and the fixed station are available. The equation reads

$$r_{\text{S3}} = \frac{1}{\text{K}_1 + \cdots + \text{K}_{\text{J}}} \sum_{j=1}^{\text{J}} \sum_{k=1}^{\text{K}_j} r(\phi_{k,j}, \lambda_{k,j}, t_{k,j}). \tag{7.15}$$

Here, J is the total number of overflights to be considered, and K_j is the total number samples which are located within the confined area of $\Omega_{2\text{km}}$ and belong to the overflight j. Figure 7.4 shows the comparison flow for this method.

$$h(\phi, \lambda, t) = h_{\text{sat}}(\phi_0, \lambda_0, t) - \rho(\phi, \lambda, t) - \Delta h_{\text{topo}}(\phi, \lambda) - C_{\text{atm}}(\phi_0, \lambda_0, t) - C_{\text{tides}}(\phi_0, \lambda_0, t) + 0$$

$$h_{\text{FS}}(t) = \bar{h}_{\text{GNSS}}^{\text{PPP}}(t) - \Delta h_{\text{offset}} - \overline{\Delta h}_{\text{MSSS}}(t) + 0$$

$$h_{\text{FS}}(\phi_{\text{FS}}, \lambda_{\text{FS}}, t_j)$$

$$h_{\text{ALT}}(\phi_{k,j}, \lambda_{k,j}, t_{k,j})$$

$$h'_{\text{DSM}}(\phi_{\text{x}}, \lambda_{\text{x}}, t_{\text{c}})$$

$$r(\phi_{k,j}, \lambda_{k,j}, t_{k,j}) = h_{\text{ALT}}(\phi_{k,j}, \lambda_{k,j}, t_{k,j}) - \left[\underbrace{h_{\text{FS}}(\phi_{\text{FS}}, \lambda_{\text{FS}}, t_j) - h'_{\text{DSM}}(\phi_{\text{FS}}, \lambda_{\text{FS}}, t_{\text{c}}) + h'_{\text{DSM}}(\phi_{j,k}, \lambda_{j,k}, t_{\text{c}})}_{h_{\text{FRM}}(\phi_{k,j}, \lambda_{k,j}, t_j)} \right] + 0$$

$$r_{\text{S3}} = \frac{1}{\text{K}_1 + \text{K}_2 + \ldots + \text{K}_J} \sum_{j=1}^{J} \sum_{\substack{k=1 \\ (\phi, \lambda) \in \Omega_{2\text{km}}}}^{\text{K}_j} r(\phi_{k,j}, \lambda_{k,j}, t_{k,j})$$

j: cycle number
J: total cycles in $\Omega_{2\text{km}}$
k: sample number
K_j: total samples of cycle j inside $\Omega_{2\text{km}}$
t_{c}: time of campaign
(ϕ, λ): POCA location
$(\phi_{\text{x}}, \lambda_{\text{x}})$: location of point x on DSM
$(\phi_{\text{FS}}, \lambda_{\text{FS}})$: representative location of fixed station

Figure 7.4. Comparison approach 1

7.5.1.2 Approach 2

We may prefer to derive the statistics with a slightly different flow. For instance, imagine a scenario where the collected heights at the fixed station are resampled to a sampling rate (t') that does not resemble that of the satellite altimetry mission (t). Or, for some reasons, there are periods of gap in the data collection at the FRM station. In either case, if the variation of topography in time is not significant, we may decide to compare the average of TDP estimates to the average of FRM estimates, as opposed to method 1, where we average the differences of TDP and FRM samples. As depicted in Figure 7.5, it is required that every single height estimate at the fixed station is first transferred to the POCA location of $(\phi_{k,j}, \lambda_{k,j})$, and then the transferred heights are averaged to represent a single FRM height, h_{FRM}:

$$h_{\text{FRM}} = \frac{1}{\text{K}_1 + \cdots + \text{K}_J} \sum_{j=1}^{J} \sum_{k=1}^{\text{K}_j} h_{\text{FS}}(\phi_{\text{FS}}, \lambda_{\text{FS}}, t'_j) - h'_{\text{DSM}}(\phi_{\text{FS}}, \lambda_{\text{FS}}, t_{\text{c}}) + h'_{\text{DSM}}(\phi_{k,j}, \lambda_{k,j}, t_{\text{c}}). \quad (7.16)$$

Same averaging is implemented for the TDP measurements,

$$h_{\text{ALT}} = \frac{1}{\text{K}_1 + \cdots + \text{K}_J} \sum_{j=1}^{J} \sum_{k=1}^{\text{K}_j} h_{\text{ALT}}(\phi_{k,j}, \lambda_{k,j}, t_{k,j}). \quad (7.17)$$

The assumption is that both h_{ALT} and h_{FRM} are representative of a height that can be associated to the confined area $\Omega_{2\mathrm{km}}$ during some specific period. Hence, an estimation of the accuracy of Sentinel-3 height measurements can be derived as

$$r_{\mathrm{S3}} = h_{\mathrm{ALT}} - h_{\mathrm{FRM}} + 0.$$

<div align="right">(7.18)</div>

$$h(\phi, \lambda, t) = h_{\mathrm{sat}}(\phi_0, \lambda_0, t) - \rho(\phi, \lambda, t) - \Delta h_{\mathrm{topo}}(\phi, \lambda) - C_{\mathrm{atm}}(\phi_0, \lambda_0, t) - C_{\mathrm{tides}}(\phi_0, \lambda_0, t) + 0$$

$$h_{\mathrm{FS}}(t') = \overline{h}_{\mathrm{GNSS}}^{\mathrm{PPP}}(t') - \Delta h_{\mathrm{offset}} - \overline{\Delta h}_{\mathrm{MSSS}}(t') + 0$$

$$h_{\mathrm{FS}}(\phi_{\mathrm{FS}}, \lambda_{\mathrm{FS}}, t'_j)$$

$$h_{\mathrm{ALT}}(\phi_{k,j}, \lambda_{k,j}, t_{k,j})$$

$$h'_{\mathrm{DSM}}(\phi_{\mathrm{x}}, \lambda_{\mathrm{x}}, t_{\mathrm{c}})$$

$$h_{\mathrm{FRM}} = \frac{1}{K_1 + \ldots + K_J} \sum_{\substack{j=1 \\ (\phi,\lambda)\in\Omega_{2\mathrm{km}}}}^{J} \sum_{k=1}^{K_j} \left(h(\phi_{\mathrm{FS}}, \lambda_{\mathrm{FS}}, t'_j) - h'_{\mathrm{DSM}}(\phi_{\mathrm{FS}}, \lambda_{\mathrm{FS}}, t_{\mathrm{c}}) + h'_{\mathrm{DSM}}(\phi_{k,j}, \lambda_{k,j}, t_{\mathrm{c}}) \right)$$

$$h_{\mathrm{ALT}} = \frac{1}{K_1 + \ldots + K_J} \sum_{\substack{j=1 \\ (\phi,\lambda)\in\Omega_{2\mathrm{km}}}}^{J} \sum_{k=1}^{K_j} h_{\mathrm{ALT}}(\phi_{k,j}, \lambda_{k,j}, t_{k,j})$$

$$r_{\mathrm{S3}} = h_{\mathrm{ALT}} - h_{\mathrm{FRM}} + 0$$

t: discrete time variable of altimeter
t': discrete time variable of fixed station
j: cycle number
J: total cycles in $\Omega_{2\mathrm{km}}$
k: sample number
K_j: total samples of cycle j inside $\Omega_{2\mathrm{km}}$
t_{c}: time of campaign
(ϕ, λ): POCA location
$(\phi_{\mathrm{x}}, \lambda_{\mathrm{x}})$: location of point x on DSM

resampling
representation
instability of FRM setup

Figure 7.5 Comparison approach 2

7.5.2   Comparison over large areas

7.5.2.1   Gridded area approach

Given the dense sampling of altimetry tracks around the poles, it is conceivable to conduct a comparison between an airborne (or ground-based) grid-line survey and the point cloud of altimetry samples within a specific area. Such a comparison may involve interpolation of the grid-lines to a continuous DSM or directly to altimetry data locations, or conversely, interpolation of the altimetry data to survey locations.

7.5.2.2   Along-track POCA approach

To avoid uncertainties from interpolation between grid-lines, an alternative is to design a survey that follows actual (at a delay) or predicted POCA ground tracks. This can be achieved using airborne campaigns with lidar or radar altimeters or GNSS receivers mounted on a sledge. However, the approach is challenging in rough or sloping terrain as the POCA track will undulate considerably. In that respect, ground-based surveys might be more flexible than airborne ones which are limited by aircraft manoeuvring. On the other hand, airborne surveys can cover a wider swath along the track than what is typically the case for ground-based surveys, so either way has pros and cons.

7.5.3   Interpreting the comparison

When there are observable differences between the satellite and the FRM, it is important to understand that these can be caused by any combination of the following:

- If the differences are smaller than the uncertainty associated with the difference (which includes components due to the satellite, the fixed station measurement and the transfer between the two using the DSM), then we can consider that the uncertainty is validating the uncertainties. Further statistical analyses can test whether the spread of individual comparisons is as expected due to the random components (which come from the measurements and from the natural variability of the surface, as discussed in Section 8).
- If the differences are larger than the uncertainty associated with the difference then it may be that:
    - The satellite has a bias that is larger than the evaluated uncertainties
    - The fixed station has a bias that is larger than the evaluated uncertainties
    - The surface is changing in a systematic way that limits the applicability of the DSM
    - Some combination of these.

If the comparison were performed without the support of a detailed understanding of the traceability and sources of uncertainty, then it would be difficult to interpret a comparison result. The analysis that has been performed and described in this report is important to understand comparison results. In the next phase of the project this analysis will be refined and quantified.

8 Annex A: The measurand for land ice

8.1 Purpose of the annex

This annex is provided to describe the nature of the true surface and to consider the impact of natural variability on the uncertainty of measurements. It supports the text in sections 7.4 and 7.5.

8.2 Notational assumptions

In the GUM, there is a notational assumption that an equation written with capitals

$$Y = f(X_1, X_2, X_3, \dots) \qquad (8.1)$$

is the quantity model, i.e., the model that describes how the quantities relate, and an equation written with small letters

$$y = f(x_1, x_2, x_3, \dots) \qquad (8.2)$$

is the model that describes the measured quantity values, i.e., what is actually measured. This notation assumes that y is an estimate of Y.

Here that notation is not used, because h is used to mean the height relative to the Earth's ellipsoid and H to mean the height relative to the Earth's geoïd. δ is also used to mean a relatively small change and Δ to mean a larger change. So, in this annex, the superscript Q is used to represent the quantity value. This could also be interpreted as the "true" value, if preferred. Other notation follows the notation used in the earlier sections.
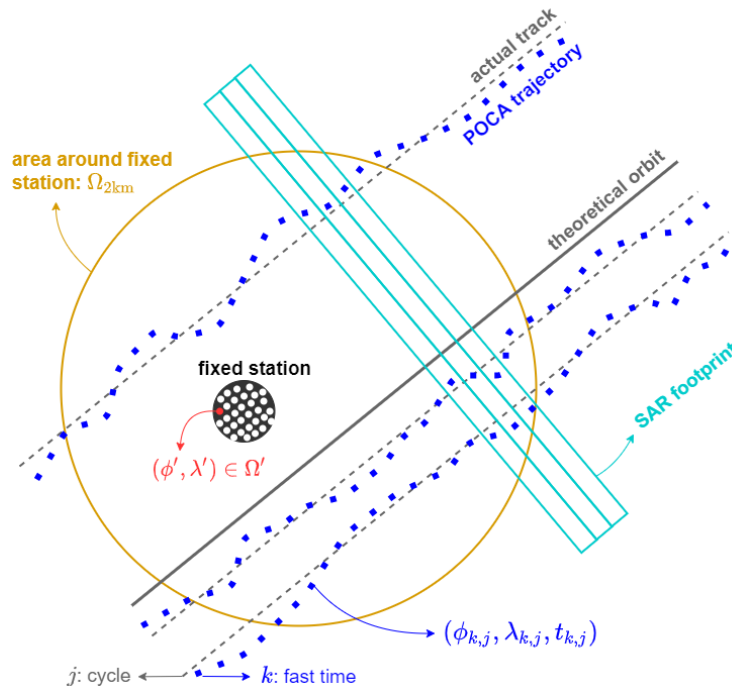
Figure 8.1 Repeat of Figure 3.4 for reference in the annex.

j describes a satellite overpass. t_j is the time of the overpass.

k describes a single POCA value as the satellite goes across the defined area around the fixed station. It is assumed that the different k are within the same time t_j.

i describes a single measurement by one of the elements of the fixed station at one time. This is at a location (ϕ'_i, λ'_i) and at time t_i.

The label subscript POCA is used to describe the satellite point of closest approach, and the label subscript FS to represent the fixed station location.

8.3 Describing the surface

The surface of an ice sheet is very complicated with multiple layers including the bed rock, the ice, the firn, and the fresh-surface now. For the purposes of a metrological analysis, however, this can be simplified into three components:

- The bedrock terrain, denoted T in equations, which changes extremely slowly (>1 year) due to historic or current loading changes and surface motion.
- The firn and more stable layer of snow, denoted F in equations, which changing slowly due to compaction and meltwater percolation, significant over weeks and months. For the purposes of this analysis, this does not strictly need to be true multiyear firn (for example, it is likely to also include some of the covered snow), but this notation is helpful to distinguish it simply from the other aspects.
- The drifting snow layer, denoted D in equations, which changes rapidly, over minutes and days.

8.4 Fixed station and satellite POCA

Figure 8.2 shows the position of the satellite POCA point, for orbit j and specific POCA point k, and the fixed station element i, for a one-dimensional slice through the surface that goes through those points. As discussed in the Section 3.1, the measured POCA is below the surface because of radar penetration of snow, however the Sentinel 3 TDP provides a POCA corrected for the surface.
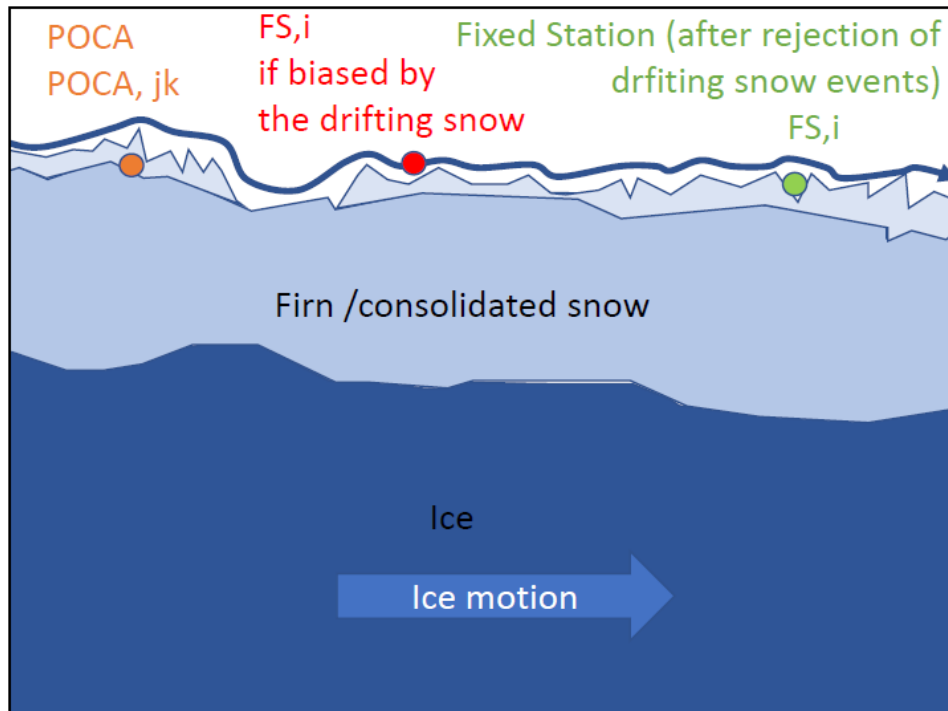
Figure 8.2 One-dimensional slice through the surface between a POCA measurement and a fixed station measurement (reused Figure 3.3)

With the defined notation, the height of the (real) surface at the POCA k at the time of the jth orbit is:

$$h_{\text{POCA},k}^{\text{Q}}(t_j) = h_{\text{T,POCA},jk}^{\text{Q}} + \delta h_{\text{F,POCA},jk}^{\text{Q}}(t_j) + \delta h_{\text{D,POCA},k}^{\text{Q}}(t_j). \qquad (8.3)$$

Note that:

- The superscript Q represents that this is a quantity equation. The satellite measured value will approximate this subject to the satellite's uncertainty. The different terms here are neither errors nor uncertainties, they relate to the natural surface variability.

- The subscript T represents the terrain. Thus $h^{\mathrm{Q}}_{\mathrm{T,POCA},jk}$ is the terrain height (above the ellipsoid) at the POCA for orbit j, kth measurement. This does not relate to a time as we assume that this is time invariant.
- The subscript F represents the firn. This is expected to change slowly, so only with the t_j time change– this is the additional thickness of the firn at the time of the overpass, but this is not expected to change much from one orbit to the next.
- The subscript D represents the windswept surface. This changes more rapidly and thus is expected to change from one orbit to the next.

Note that the measurements may not be able to distinguish these separate elements and they may not correspond perfectly to physical definitions. These distinctions are added to the equations to explain the extent to which natural variability can be smoothed or averaged out by multiple measurements, and therefore they are theoretical concepts. Similarly, the superscript Q shows that here 'true' quantity values are considered, and no measurement uncertainty is included. Any real measurement will also include measurement uncertainty. This is discussed in Section 8.9.

Similarly, the height of the real surface at the fixed station's ith element at a time t_j is:

$$h^{\mathrm{Q}}_{\mathrm{FS},i}(t_j) = h^{\mathrm{Q}}_{\mathrm{T,FS},i} + \delta h^{\mathrm{Q}}_{\mathrm{F,FS},i}(t_j) + \delta h^{\mathrm{Q}}_{\mathrm{D,FS},i}(t_j). \qquad (8.4)$$

8.5 Transforming the fixed station quantity to the POCA to obtain the FRM

A comparison of the fixed station measured value with the satellite measured value, requires the relative height difference:

$$\Delta^{\mathrm{Q}}_{\mathrm{FS\text{-}POCA},ijk}(t_j) = h^{\mathrm{Q}}_{\mathrm{FS},i}(t_j) - h^{\mathrm{Q}}_{\mathrm{POCA},k}(t_j). \qquad (8.5)$$

This is evaluated using the DSM determined during the field campaign. The field campaign took place at a time t_{C}, where the subscript label C represents "campaign".

At the time of the field campaign, the real surface height is:

$$h^{\mathrm{Q}}_{\mathrm{POCA},k}(t_{\mathrm{C}}) = h^{\mathrm{Q}}_{\mathrm{T,POCA},jk} + \delta h^{\mathrm{Q}}_{\mathrm{F,POCA},jk}(t_{\mathrm{C}}) + \delta h^{\mathrm{Q}}_{\mathrm{D,POCA},k}(t_{\mathrm{C}}) \qquad (8.6)$$

at the POCA position that will become k of orbit j (i.e., this is not at the time of that overpass, but it's at the location where that overpass will occur in the future).

Similarly, at the fixed station, the real surface height at the time of the campaign is:

$$h^{\mathrm{Q}}_{\mathrm{FS},i}(t_{\mathrm{C}}) = h^{\mathrm{Q}}_{\mathrm{T,FS},i} + \delta h^{\mathrm{Q}}_{\mathrm{F,FS},i}(t_{\mathrm{C}}) + \delta h^{\mathrm{Q}}_{\mathrm{D,FS},i}(t_{\mathrm{C}}). \qquad (8.7)$$

Because the DSM was only determined at the time of the field campaign, the only knowledge of the height difference is:

$$\Delta^{\mathrm{Q}}_{\mathrm{FS\text{-}POCA},ijk}(t_{\mathrm{C}}) = h^{\mathrm{Q}}_{\mathrm{FS},i}(t_{\mathrm{C}}) - h^{\mathrm{Q}}_{\mathrm{POCA},k}(t_{\mathrm{C}}). \qquad (8.8)$$

Comparing equations (8.5) and (8.8), while inserting equations (8.3), (8.4), (8.6) and (8.7), shows that the terrain part cancels out (because, by definition, that is unchanging between the campaign and the overpass).

In other words, subtracting the known height difference from the campaign $\Delta^{\mathrm{Q}}_{\mathrm{FS\text{-}POCA},ijk}(t_{\mathrm{C}})$, from the true height difference at the time of the overpass $\Delta^{\mathrm{Q}}_{\mathrm{FS\text{-}POCA},ijk}(t_{\mathrm{C}})$ gives:

$$\begin{aligned}
&\Delta^{\mathrm{Q}}_{\mathrm{FS\text{-}POCA},ijk}(t_j) - \Delta^{\mathrm{Q}}_{\mathrm{FS\text{-}POCA},ijk}(t_{\mathrm{C}}) \\
&= h^{\mathrm{Q}}_{\mathrm{FS},i}(t_{\mathrm{C}}) - h^{\mathrm{Q}}_{\mathrm{POCA},k}(t_{\mathrm{C}}) - h^{\mathrm{Q}}_{\mathrm{FS},i}(t_j) + h^{\mathrm{Q}}_{\mathrm{POCA},k}(t_j)
\end{aligned} \qquad (8.9)$$

$$\begin{aligned}
&= h_{\text{T,FS},i}^{\text{Q}} + \delta h_{\text{F,FS},i}^{\text{Q}}(t_{\text{C}}) + \delta h_{\text{D,FS},i}^{\text{Q}}(t_{\text{C}}) - h_{\text{T,POCA},jk}^{\text{Q}} - \delta h_{\text{F,POCA},jk}^{\text{Q}}(t_{\text{C}}) \\
&\qquad - \delta h_{\text{D,POCA},k}^{\text{Q}}(t_{\text{C}}) - h_{\text{T,FS},i}^{\text{Q}} - \delta h_{\text{F,FS},i}^{\text{Q}}(t_j) - \delta h_{\text{D,FS},i}^{\text{Q}}(t_j) \\
&\qquad + h_{\text{T,POCA},jk}^{\text{Q}} + \delta h_{\text{F,POCA},jk}^{\text{Q}}(t_j) + \delta h_{\text{D,POCA},k}^{\text{Q}}(t_j) \\
&= \left(\delta h_{\text{F,FS},i}^{\text{Q}}(t_{\text{C}}) - \delta h_{\text{F,FS},i}^{\text{Q}}(t_j) \right) - \left(\delta h_{\text{F,POCA},jk}^{\text{Q}}(t_{\text{C}}) - \delta h_{\text{F,POCA},jk}^{\text{Q}}(t_j) \right) + \\
&\qquad \left(\delta h_{\text{D,FS},i}^{\text{Q}}(t_{\text{C}}) - \delta h_{\text{D,FS},i}^{\text{Q}}(t_j) \right) - \left(\delta h_{\text{D,POCA},k}^{\text{Q}}(t_{\text{C}}) - \delta h_{\text{D,POCA},k}^{\text{Q}}(t_j) \right).
\end{aligned}$$

Here, the equation is colour-coded to make the different components easier to identify.

Note, again, that these equations do not describe the measurements, nor include any measurement uncertainty. This simply describes the natural variability of the surface between the time of the campaign and the time of the satellite overpass.

8.6 Reducing the drifting snow / wind-swept term by averaging

The red terms in Equation (8.9), those with the subscript D, represent the relative thickness difference of the drifting snow (rapidly varying) part of the surface between the fixed station and the POCA position, and between the time of the campaign and the time of the overpass.

For any one measurement this cannot be known because, by definition, this changes more regularly than the observations. However, that rapid, and random variation means that the drifting part of the surface is changing rapidly in time and over small spatial scales, and therefore its impact can be reduced by averaging.

8.6.1 Averaging the fixed station measurements

For the fixed station, the averaging is explicit. There are several different fixed station positions, separated by a few metres, with the number ranging from about 7 (for the ultrasonic stations) to around 36 (for the MSSS). Additionally, a temporal average over say 1 day of measurements around the time of the overpass can provide a further average.

In Section 7.4, the average of the fixed station readings is given as

$$\overline{\Delta h}_{\text{MSSS}}(\tau) = \underset{(\phi',\lambda') \in \Omega', \tau' \in \left[\tau - \frac{\delta\tau}{2}, \tau + \frac{\delta\tau}{2}\right]}{\text{mean}} \Delta h_{\text{MSSS}}(\phi', \lambda', \tau')$$

Repeat of eq. (7.9)

This is subtly different from averaging equation (8.7), because in the real measurements, the fixed station readings are averaged and then subtracted from the GNSS position, which is measured at different times. However, in this annex the emphasis is on the true quantities of the surface height in order to distinguish natural variability that is independent of the method of measurement.

8.6.2 For the satellite POCA

At the satellite POCA for any one overpass, for any one measurement value k, there is only one measurement and therefore there is no direct averaging out of this variability. However, a POCA elevation is determined at a coarser resolution than individual surface rangers at a station and thus we can assume that the representative spatial scale is more similar to a multi-instrument station average. Furthermore, any residual spatial sampling error can be considered unbiased. That is, at any one location, it is drawn from a symmetrical distribution. Therefore, over multiple satellite measured values at different POCAs across the area Ω', the effect cancels out.

8.7 Reducing the uncertainty associated with the longer-term changes

The situation with the slowly varying part, labelled F for firn, but also including the part of the snow not blown about rapidly (the blue parts of Equation (8.9)), is more complex. From a mathematical perspective, in order to use the DSM determined at the time of the campaign, one of the following must be true:

a) the surface change is small, i.e., $\left(\delta h_{\text{F,FS},i}^{\text{Q}}(t_j) - \delta h_{\text{F,FS},i}^{\text{Q}}(t_{\text{C}}) \right)$ and $\left(\delta h_{\text{F,POCA},jk}^{\text{Q}}(t_j) - \delta h_{\text{F,POCA},jk}^{\text{Q}}(t_{\text{C}}) \right)$ are both small and can be considered negligible

b) there is a systematic relationship between the firn change at the point of overpass and the firn change at the fixed station. This means that $\left(\delta h_{\mathrm{F,FS},i}^{\mathrm{Q}}(t_j) - \delta h_{\mathrm{F,FS},i}^{\mathrm{Q}}(t_{\mathrm{C}})\right)$ is equal to $\left(\delta h_{\mathrm{F,POCA},jk}^{\mathrm{Q}}(t_j) - \delta h_{\mathrm{F,POCA},jk}^{\mathrm{Q}}(t_{\mathrm{C}})\right)$, so that the difference between them is negligible

c) there is a difference between the two changes that varies randomly across the site, and therefore averages out for many overpasses.

d) There is a difference between the two changes that varies systematically across the site (e.g., with a slope), but there are sufficiently equal numbers of measurements on one side of that slope as on the other, that the problem is symmetrical (the low ones cancel the high ones).

In the practical situation, immediately after the campaign, and before any snow fall, an assumption of situation (a) is reasonable: at that point the campaign DSM is basically valid (apart from the wind-blown snow drift that changes rapidly and is reduced by averaging). An uncertainty due to natural variability can be included to account for the extent this is not negligible and will increase as the time since the measurement campaign has increased, especially following precipitation.

The time series of measurements at the fixed station provides a correction for any semi-systematic relationship, as described in situation (b). Of course, precipitation across the site will not be even, and is affected by the position of the sastrugis. The sastrugis themselves, and their spatial movement, create some semi-random variability across site, as in situation (c). Situation (d) applies if there is some systematic slope to the snow fall across the site, perhaps because of the underlying terrain shape or the shadowing effect of nearby terrain.

In general, the assumption is made that the DSM at the time of the campaign is valid to correct fixed station measurements for the following year. As discussed here, there is an uncertainty associated with this assumption that will need to be considered in the analysis.

8.8   Quantifying and validating the effect of natural variability

The discussions in sections 8.6 and 8.7 considered that the different sources of natural variability are expected to cancel out in the comparisons, if enough are averaged. Analysis by Libois et al. of natural variability shows how daily variations are often local, and sometimes large, even larger than the annual accumulation, as seen in Figure 5.5. From their analysis, they determined that snow transport tends to level out the snow surface over the long term, which justifies the averaging of satellite data over a 2 km circle.

8.9   Measurement uncertainties

In this analysis, only natural variability has been considered. Measurement uncertainty was not considered in the earlier sections of this annex and all the measurement models were described in terms of the quantity (true value), rather than the measured value. However, all measurements are also sensitive to measurement uncertainty.

If there are random effects affecting the fixed station measurements, then these will be indistinguishable from the random variation of the windswept part of the surface. The impact of the saltation layer is almost random, but with a positive bias.

Random effects affecting the satellite measurements will create variations in the difference k values. These will be indistinguishable from natural variability of the POCA surface, whether due to the windswept surface or spatial variability of the firn.

Systematic affects in the fixed station measurement and / or in the satellite measurement will cause a bias in the comparisons of the satellite to the FRM. This would become apparent in step (iii) of the analysis described in section 8.8. In analysis, it is important to evaluate the uncertainties associated with the fixed station measurement and with the satellite, so that it can be determined whether the uncertainty associated with the comparison of the two is bigger or smaller than the observed bias.

Partially systematic affects will be most difficult to quantify. For example, the fixed station measurements have an uncertainty component that is sensitive to air temperature, and that changes systematically over time (due to seasons), then this will create a trend that may be indistinguishable from movements of sastrugis.