The Weather Forecast Processing Chain

Methodology and Technical Documentation

topo^alpin

Paper version 1.2 — 14 November 2026
Describes the forecast chain of topo^alpin 5.9.4

Abstract. This document specifies the deterministic multi-model post-processing chain with which topo^alpin produces site-specific mountain weather guidance for the Alps. Forecast fields from six numerical and data-driven prediction systems are weighted, combined and transformed into estimates of temperature, wind, gusts, freezing level, rain–snow line, precipitation phase, new-snow depth, thunderstorm risk and a weather symbol for named locations and their elevation bands. The chain comprises availability handling per model, elevation adjustment, pressure-level interpolation, wet-bulb-based phase partitioning, a temperature-index snow-to-liquid ratio and rule-based symbol generation. Model weights and several thresholds and correction factors are prescribed on physical or operational grounds and have not been calibrated against independent observations. The figures reported here come from internal consistency checks and sensitivity analyses on recorded model output; they are not a verification of forecast accuracy. Section 16 sets out the resulting limitations one by one.

Contents

1 Introduction 5
1.1 Decision-Relevant Variables for Mountain Tours 5
2 State of the Art 6
2.1 Snow-to-Liquid Ratio 6
2.2 Precipitation Phase 6
2.3 Gust Parameterisation 7
2.4 Multi-Model Combination 7
3 Data Basis 8
3.1 Data Access 8
3.2 Models 8
3.3 Derivation of Undocumented Input Variables 9
3.4 Temporal Structure 9
3.5 Vertical Reference 10
3.6 Vertical Profile 11
4 Multi-Model Blend and Weighting 13
4.1 Base Weights 14
4.2 Lead-Time-Dependent Reweighting 14
4.3 Aggregation Rules 14
4.4 Model Families 15
5 Freezing Level 16
5.1 Method 16
5.2 Validity Conditions 16
5.3 Interpretation of the Displayed Row 17
5.4 Rain–Snow Line 18
5.5 Snowfall Level in ICON 20
5.6 Wet-Bulb Temperature at the Site 21
6 Precipitation Phase 23
6.1 A Site-Specific Classification 23
6.2 Partitioning at Site Elevation 24
6.3 No Additional Temperature Threshold 24
7 Snow-to-Liquid Ratio 25
7.1 Scope of Application 25
7.2 Temperature Dependence 26
7.3 Wind Correction 27
7.4 Why No Damping Is Applied at the Freezing Level 27
7.5 Comparison with the Profile-Based Method 28
8 Wind 28
8.1 Vertical Assignment and Blending 28
8.2 Block Aggregation 29
8.3 Exposure Adjustment 29
8.4 Gusts 29
9 Lifting Condensation Level 30
9.1 Calculation 30
9.2 What the Variable Does Not Describe 31
9.3 Application 31
10 Thunderstorm Risk 31

10.1 Composition ... 31
10.2 Adjusted CAPE Thresholds ... 33
10.3 Humidity Condition ... 33
10.4 Cloud Layers ... 34
10.5 Achievable Score ... 34
10.6 Status of the Score ... 34
11 Weather Symbol ... 35
11.1 Severity Ranking ... 35
11.2 Voting ... 35
11.3 Correction Cascade ... 36
12 Further Variables ... 38
12.1 Sunshine Duration ... 38
12.2 Wind Chill ... 40
12.3 Precipitation Type ... 40
12.4 Other Blended Values ... 40
12.5 Computed but Not Displayed ... 42
13 Elevation Bands ... 43
13.1 Phase per Elevation Band ... 43
13.2 Which Bands Are Shown ... 44
14 Marking of Estimated Values ... 46
14.1 Why the Snow Amount Carries No Marking ... 47
14.2 Water Phase and Precipitation Type ... 47
14.3 Symbol and Amount Answer Different Questions ... 48
14.4 When a Dash Replaces a Number ... 48
15 Prescribed Parameters ... 49
16 Limitations and Open Issues ... 51
16.1 Point Accuracy versus Model Resolution ... 51
16.2 Undocumented Input Variables ... 51
16.3 The Snow-to-Liquid Ratio Remains a Heuristic ... 52
16.4 Limits of Direct Model Gusts and Model Spread ... 52
16.5 Uncalibrated Parameters ... 54
16.6 Different Limit Variables in ICON ... 55
16.7 Ranking of Fog ... 55
16.8 Different Model Sets for Score and Symbol ... 55
16.9 Limit of Phase Determination ... 55
16.10 Vertical Reference of the Phase Partitioning ... 55
16.11 Weather-Independent Elevation Correction for Three Models ... 56
16.12 Estimates from the Site ... 57
16.13 Display Choices Without Validation ... 57
16.14 Changing Contributor Set of the Probability ... 58
16.15 The Forecast Precipitation Amount ... 58
16.16 The Layer Above the Site ... 59
16.17 Unit Handling in the Chain ... 59
16.18 Lightning Potential and Updraft Without a Threshold ... 59
17 Worked Examples ... 59
17.1 Gust Without a Model Value ... 59
17.2 New-Snow Amount ... 60
17.3 Phase at the Site ... 60
18 Reproducibility ... 61

18.1 Recorded Interface Responses 61
18.2 Live Queries 61
References 61

1 Introduction

Weather forecasts for the Alpine region face a resolution problem: the model grids used here have a horizontal spacing between about 2 km and 28 km, whereas a route-planning decision in mountain terrain is made for an individual summit or ridge. Between the model output and a number that applies to a named point there is therefore a chain of aggregation, conversion and correction steps.

The models themselves are documented — AROME (Seity et al. [1]), ICON (Zängl et al. [2]) and ECMWF's data-driven AIFS (Lang et al.[3]). Their relative skill has been examined in the DIMOSIC comparison from identical initial conditions (Magnusson et al. [4]). The downstream processing, by contrast — weighting, phase determination, snow-amount estimation, symbol generation — is application-specific and generally not disclosed.

This document describes how topo^alpin — an application for tour planning, navigation and mountain weather in the Alps (web and iOS) — implements this chain. It serves the same purpose as the methodology papers on avalanche terrain and on Tour-Check. Where methods are adopted from the literature, the source is given; our own adaptations and prescribed parameters are identified as such and summarised in Section 16.

One limitation applies to the entire document and is stated at the outset: **none of the methods described here is calibrated against observations.** The prescribed parameters — model weights, gust and exposure factors, CAPE thresholds, the breakpoints of the snow-amount estimate — are justified on physical and meteorological grounds, but not fitted to measurements. Section 16.5 lists them in full; the prerequisites for a validation are described, the validation itself is outstanding.

The numbers quoted throughout — hours, blocks, medians, counts — come from three kinds of internal analysis, none of which involves independent observations: **regression tests** of the implementation against recorded interface responses, **consistency checks** comparing alternative fields of the same models, and **sensitivity analyses** in which one input or assumption is changed. Where the text speaks of an error, a reference value is named; the words skill and accuracy are reserved for comparisons against observations.

The physics, data assimilation and grid structure of the weather models are not the subject of this document. They are documented in the model descriptions (Seity et al.[1], Zängl et al.[2], Lang et al.[3]).

1.1 Decision-Relevant Variables for Mountain Tours

The variables relevant to route planning are determined by exposure and consequence rather than by meteorological completeness. Wind and gusts determine whether a ridge can still be traversed with control; temperature and wind together determine the thermal load. The freezing level, precipitation phase and new-snow amount determine whether rock, paths and approaches are snow-free, wet, snow-covered or icy. Convective indicators primarily inform timing rather than the decision to undertake a route — they shift start times and argue against exposed objectives.

The chain described here is oriented towards these variables. This explains why the following sections deal with gusts, the snow-to-liquid ratio, the freezing level and thunderstorm risk rather than with a general weather situation: a daily value for a region answers none of the questions above.

What the chain provides is a translation — from model fields computed for grid cells and standard levels into a representation for a named point and its elevation bands. What it does not provide is a go-ahead: terrain, route, group and observation on the spot lie outside what a weather model can represent.

2 State of the Art

2.1 Snow-to-Liquid Ratio

Precipitation is converted into new-snow depth via the snow-to-liquid ratio (SLR). Four approaches are in common use.

The **fixed 10:1 assumption** is widespread but demonstrably inadequate: Roebber et al.[5] show that snow density depends on seven factors, including temperature in several layers, relative humidity and external compaction by wind.

Temperature-based methods derive the SLR from a single temperature variable, usually the maximum of the air column below 500 hPa. They are simple to compute and do not require a vertical profile.

Profile-based methods weight layer-wise SLR values over the temperature profile. Cobb and Waldstreicher [6] present such a method; it underlies the operational implementation of the US National Weather Service, which uses the method's 2011 layer table; its product description of 2013 lists that table [7]. The table reaches its maximum at −16 °C and decreases again towards lower temperatures.

Statistical and machine-learning methods model the SLR as a function of several predictors. Rosenow et al.[8] compare operationally used methods across five winter storms.

Two recent studies train random-forest models on manual measurements of new snow and water equivalent. Veals et al.[9] use fourteen mountain stations in the western United States and achieve a coefficient of determination of 0.43 with a mean absolute error of 2.94; in the same comparison the operationally used methods lie between 0.04 and 0.23 and between 4.01 and 9.45 respectively. Pletcher et al.[10] repeat this across the contiguous United States, using reports from a volunteer network and reanalysis profiles, with a mean absolute error of 2.92 compared with 3.51 to 4.76.

Two aspects of this are notable for the application described here: the dominant predictors are near-surface variables — temperature, wind speed, humidity — and the margin over the operational curves is substantial.

Approach	Input variables	Vertical profile required	Reference
Fixed ratio	–	No	–
Temperature-based	Column maximum	No	–
Profile-based	Layer-wise T, humidity, vertical velocity	Yes	[6], [7]
Statistical / machine-learning	Several predictors	partly	[5], [9], [10]

Table 1: Approaches to determining the snow-to-liquid ratio.

topo^alpin uses a temperature-based method, since the Open-Meteo interface does not provide a complete vertical profile for every model and time step. The deviation from the profile-based method is quantified in Section 7.5.

2.2 Precipitation Phase

A universal 0 °C dry-bulb threshold is not generally optimal for separating rain and snow and does not represent the observed spatial variability of the transition. Jennings et al.[11] analysed 17.8 million

observations over 29 years in the Northern Hemisphere: the temperature at which rain and snow are equally frequent is 1.0 °C on average and lies between −0.4 °C and 2.4 °C for 95 % of stations.

Methods based on the **wet-bulb temperature** perform better, since it is closer to the surface temperature of a falling hydrometeor and takes humidity into account.

How much can be gained by this is constrained by a hemispheric analysis. Fontrodona-Bach et al.[12] test, using station data from 1950 to 2023, how well a simple temperature-index snow model with thresholds for accumulation and melt performs. The subject is therefore the skill of that model, not phase determination as such — the threshold enters the model as only one of several variables. Three results are relevant to the phase question: the 0 °C threshold captures 89 % of all days with snow accumulation — for melt days the same figure is 76 %, which is not pursued further here because the method computes no melt; on a daily-mean basis no fixed threshold separates rain and snow optimally, because low values classify snow days as rain and higher values classify rain as snow; and a uniform large-scale shift of the thresholds does not consistently improve model skill, but shifts it in either direction depending on the target variable.

For the Alps, the authors also note that precisely there — under maritime influence and with a deep snowpack — snow accumulation at temperatures just above freezing is difficult to capture, and they explicitly recommend a different phase method for such conditions, for instance one based on the wet-bulb temperature. The gain thus lies in the marginal range, not across the board.

Jennings et al.[13] tested three machine-learning methods against strong benchmarks on two large datasets and found only negligible additional skill — at most 0.6 % overall — when only near-surface meteorological predictors were available; the difficulty concentrates between about 0 and 4 °C. This finding limits what can be expected from refining a near-surface method.

2.3 Gust Parameterisation

Brasseur[14] derives gusts physically from the deflection of air parcels brought down by turbulent eddies from higher layers of the boundary layer. The method requires a profile of wind speed and turbulent kinetic energy.

The method presupposes boundary-layer variables that the interface used does not provide. topo^alpin therefore uses an empirical elevation-dependent factor (Section 8.4).

2.4 Multi-Model Combination

Several models can be combined with equal or unequal weights. A robust weighting requires that model skill can be separated from differences in the initial data. Magnusson et al.[4] achieve this in the DIMOSIC project by running seven global models over 122 dates from identical initial conditions. In this set-up the IFS shows the smallest root-mean-square error for most variables and regions, with ICON in second place.

The main conclusion of the study, however, qualifies any weighting by model formulation: from the same initial data all seven models perform similarly well, which is why the authors identify the initial conditions as the stronger factor behind differences in skill between forecasting centres. In operational practice, where each centre runs its own assimilation, the difference persists — but what can be derived from DIMOSIC is only the ranking among **global** models, not the skill of a regional model.

3 Data Basis

3.1 Data Access

All model data are obtained via Open-Meteo[15]; the application does not query any weather service directly. A dedicated Open-Meteo instance is operated for this purpose.

3.2 Models

Model	Operator	Grid used	Native mesh	Forecast range
ICON D2 [2]	DWD	0.02° (~2 km)	2.2 km	~ +47 h
AROME France [1]	Météo-France	0.025° (~2.5 km)	1.3 km	~ +47 h
ICON EU [2]	DWD	0.0625° (~7 km)	6.5 km	~ +118 h
ECMWF IFS	ECMWF	0.25° (~25 km)	9 km	~ +230 h
AIFS [3]	ECMWF	0.25° (~28 km)	–	~ +230 h
GFS	NOAA	0.11° (~13 km), pressure levels 0.25°	13 km	~ +230 h

Table 2: Models used. *Grid used* is the product grid that the application ingests through the interface, as stated in the Open-Meteo documentation (accessed 9 October 2026); *native mesh* is the operator's computational grid. For IFS and AROME the two differ substantially. The forecast ranges are those observed at the interface and vary over the run cycle by about ± 6 h.

AIFS is ECMWF's data-driven forecasting system (Lang et al.[3]). It has been running operationally alongside the physics-based IFS since 25 February 2025 — ECMWF's first operational machine-learning forecasting system[16].

Variable	Available in
Precipitation, temperature, 10-m wind, humidity	all
Surface pressure	all
Gusts	all except AIFS
Upper-level wind 850/700 hPa	all
Pressure-level temperature, humidity and geopotential height	all, on different levels per model (Section 3.6)
Snowfall and rain amounts separately	all
Snow water equivalent	all
Wet-bulb freezing level	ICON EU, ICON D2
WMO weather code [17]	all
Precipitation probability	ECMWF, GFS, ICON EU
Precipitation type	ECMWF only
Sunshine duration	all
Snow depth	ICON EU, ICON D2, ECMWF, GFS
CAPE	all except AIFS
Showers (convective precipitation)	all
Convective inhibition, lifted index	GFS only
Convective cloud top	ICON EU, ICON D2
Boundary-layer height	GFS only

Variable	Available in
Total cloud cover	all
Low, mid and high cloud cover	all

Table 3: Availability of the **evaluated** variables per model, measured at the interface across all six models. Missing values are excluded from the averaging, not counted as zero.

The request also retrieves variables that are **not yet evaluated**. They are listed separately here because their presence would otherwise be read as use: they are raw material for named open issues, not part of any computation.

Variable	Available in
Provider's wet-bulb temperature	all
Convective cloud base	ICON EU, ICON D2
Lightning potential	ICON D2 only
Updraft	ICON D2 only
Wind at 80–200 m	all
Vapour pressure deficit	all
Surface temperature	all except AROME
Direct normal irradiance	all
Total column water vapour	ECMWF, GFS
UV index	GFS only
Runoff	ECMWF only

Table 4: Variables retrieved but not evaluated. Availability is measured on the same responses as above. Their intended use is stated in Section 16.

Low and high cloud cover are available for all models but are evaluated only from ICON — a decision, not a lack of availability (Section 10.4). That convective inhibition and the lifted index come exclusively from GFS is decisive for the thunderstorm risk score (Section 10.1, Section 16.8).

3.3 Derivation of Undocumented Input Variables

CAPE, convective inhibition, lifted index, freezing level, snowfall and apparent temperature are not supplied by the models but derived by the data provider. The documentation gives no references for these derivations; it cites one source for cloud cover, refers to FAO-56 for evapotranspiration and to WMO guidelines for the UV index. For the apparent temperature, the scale after Steadman[18] is given; the value is adopted unchanged.

Our own processing is thus traceable to sources, the derivation of these input variables is not. The consequences are described in Section 16.2.

3.4 Temporal Structure

Ten forecast days and three past days are retrieved at hourly resolution. The evaluation combines them into four six-hour blocks per day (00–06, 06–12, 12–18, 18–24 local time); for the short-range view the same blocks exist with three-hour and one-hour length. All three resolutions pass through the same processing chain.

Open-Meteo stamps each sum — precipitation, sunshine duration, the gust as a maximum — at the end of its hour: the value at 13:00 covers 12 to 13. A block therefore takes the stamps after its start

up to its end, 12–18 the stamps 13 to 18, and the state variables are read at the same stamps. The evening block closes with midnight of the following day; for the last day, the request reaches one day further than the display. A one-hour column carries the stamp it is labelled with. On the days the clock changes, a block therefore holds seven or five stamps. Hours are tracked by their position in the response, not by date and clock time, which repeat in autumn; date and time serve only as labels.

Within a block, amounts are summed (precipitation, snow, rain, sunshine duration), state variables are averaged (wind, humidity, condensation level), extreme variables are carried as maximum or minimum (temperature, gusts, CAPE, apparent temperature), and the weather code is read from the middle of the block. This applies to the hours **within** a block; across models the gust follows a rule of its own (Section 8.4). The total snow depth is a state value and is taken from the last hourly value of the block.

The weather code from the middle of the block is a **starting value**, not the last word. In stable weather it provides a representative state and prevents a single short shower from depicting an entire block as rain. But it describes one hour out of six, and two corrections are applied afterwards.

If a more severe state persists long enough within the block, it takes over (Section 11.2). Otherwise a block with five hours of showers and a dry middle hour would carry a cloud symbol.

If the code contradicts the precipitation amount of the block, the correction cascade applies (Section 11.3). A dry mid-block code next to a visible amount would be correctly aggregated in itself, but contradictory in the cell, because the display does not explain the different reference times.

Thresholds that act on amounts — the correction cascade of the weather symbol (Section 11.3) — scale with block length, so that a one-hour block is not measured against a six-hour threshold.

3.5 Vertical Reference

Three elevations must be distinguished.

The **site elevation** is the elevation of the named point, taken from OpenStreetMap or otherwise from our own terrain model. It is passed to the interface, which then adjusts the near-surface fields to this elevation. The nature of this adjustment is **model-dependent** and not documented by the data provider (Section 3.3); it can, however, be measured by querying the same coordinate twice with different elevations.

Over 48 hours at the Zugspitze, `elevation=1000` versus `elevation=3000`:

Model	ΔT median	ΔT range	K per 100 m
ICON EU	−10.25	−12.00 … −8.40	0.51
ICON D2	−13.10	−15.00 … −9.50	0.66
AROME	−13.10	−14.30 … −11.90	0.66
ECMWF IFS	−13.00	−13.00 … −13.00	**0.650**
GFS	−13.00	−13.00 … −13.00	**0.650**
AIFS	−13.00	−13.00 … −13.00	**0.650**

Table 5: Effect of the elevation parameter on the 2-m temperature, measured from two queries of the same coordinate and the same period.

For ECMWF, GFS and AIFS it follows a fixed rate of 0.650 K per 100 m over the period examined — in every single hour, without scatter. For ICON EU, ICON D2 and AROME it varies with time;

ICON EU, at 0.51 K per 100 m, also lies systematically away from the standard rate. A second indication of the same split is relative humidity: in the first group temperature and dew point shift by the same amount, so that it remains unchanged across both elevations; for ICON D2 it moves from 79 to 85 %, for AROME from 86 to 97 %.

The processing adopts the values provided at site elevation and does not apply a second temperature–elevation correction. What follows from this for phase determination is set out in Section 16.11.

The site elevation also controls the choice of wind source (Section 8.1), the gust factor (Section 8.4) and — via the wet-bulb temperature — the partitioning of precipitation into a solid and a liquid fraction (Section 6.2).

The lifting condensation level is computed as height above ground from temperature and dew point at site elevation, and it is converted to elevation above sea level with the same site elevation (Section 9.1).

If neither OpenStreetMap nor the terrain model supplies a site elevation, the grid-point elevation takes its place. For a query covering several models, the interface returns only one elevation value — that of the first model — so that in this fallback the elevation of one model controls the processing of all six. The case is named in the data model and thus remains distinguishable.

The spread between these elevations is considerable and is assessed in Section 16.1.

3.6 Vertical Profile

Several variables require the state of the atmosphere at an elevation other than the site elevation: the freezing level, the rain–snow line and the phase for each elevation band of the display grid. They are based on a profile built from the pressure levels of the respective model.

Temperature, relative humidity and geopotential height are requested on six levels. Which of them a model provides differs:

Level	ICON EU	ICON D2	AROME	ECMWF	GFS	AIFS
950 hPa, ≈ 540 m	yes	yes	yes	–	yes	–
925 hPa, ≈ 760 m	yes	–	yes	yes	yes	yes
850 hPa, ≈ 1492 m	yes	yes	yes	yes	yes	yes
800 hPa, ≈ 2030 m	yes	yes	yes	–	yes	–
700 hPa, ≈ 3097 m	yes	yes	yes	yes	yes	yes
600 hPa, ≈ 4200 m	yes	yes	yes	yes	yes	yes
500 hPa, ≈ 5742 m	yes	yes	yes	yes	yes	yes

Table 6: Pressure levels per model. The heights are median values from 6940 model hours over three Alpine summits; they vary with the weather situation by several hundred metres.

The selection follows the elevations that the display grid actually shows, not an even spacing. With 850, 700 and 500 hPa alone, the 1000-m row would lie below the lowest level in **every one** of the 6940 measured model hours, and between 700 and 500 hPa there would be a gap of 2645 m. 925 and 950 hPa cover the lower row — two levels, because no single level is available in all five models —, 800 hPa carries the touring band, and 600 hPa halves the upper gap and secures the 4000-m row.

800 hPa at around 2030 m closes the gap in the touring band. Without this level, 700 hPa at 3097 m would be the lowest above the orography, and every summit below it would be left without a profile phase — even though the model computes there. Measured at twelve Alpine sites over 360 hours each in January and July, it carries the summit row at the Zugspitze with 0.60 instead of 0.30 of the

model weight, at the Nebelhorn with 1.00 instead of 0.50, at Triglav with 0.80 instead of 0.40. The lower rows do not depend on it, and high summits such as Mont Blanc carry their row in any case.

900 hPa is not used. It lies at around 990 m — below almost all model terrain in the Alpine region and below every row drawn — and in the same 24 cases did not yield a single additional value.

3.6.1 Derived Pressure Levels

The data service issues levels that the model itself does not provide. ICON D2 delivers 925 hPa although its level ladder jumps directly from 850 to 950: measured over 132 hours, the relative humidity there is **exactly** the linear blend of the two neighbours, with a deviation of 0.0000 %. For ICON EU, which genuinely provides the level, the deviation is up to 0.96 %, for GFS 10.1 %.

The detection reaches a limit, however, and 900 hPa shows it: for ICON D2 the humidity there deviates by at most 0.6 percentage points from the blend of 950 and 850 hPa over 4320 hours examined. That is practically the interpolation itself, but not **exactly** — a test for vanishing deviation does not flag it. A tolerance threshold would be the wrong way out: in well-mixed air, humidity genuinely varies almost linearly, and the same level would then be native or not depending on the weather situation. A provenance that changes with the weather is no provenance.

The provenance is therefore **fixed**, once, when a level is added, and tested against the pair of neighbours that the domain provides *without* it. This is essential: if 900 hPa were in the ladder, 925 would suddenly be tested against 950 and 900 instead of against 950 and 850 — and the case being sought would disappear behind the new intermediate level. Detection from the values remains in place alongside; it catches only exact interpolation, so it cannot wrongly reject anything, and serves as a safety net for levels about which nothing has been fixed.

Such levels are removed from the profile. Inserted into a linear interpolation they would change nothing — they lie on the same straight line, after all — but they would feign a vertical resolution that the model does not have, and this pretence would show up in the apparent agreement of the models. The test runs over the entire time series: a single hour can lie on the line by chance, hundreds in succession cannot.

3.6.2 Interpolation to an Intermediate Elevation

Between two levels, temperature is interpolated linearly in height and pressure log-linearly between the two pressure coordinates — this suits pressure levels and is more consistent than extrapolating barometrically upwards from surface pressure.

Water vapour is **not** interpolated as relative humidity but as mixing ratio. From T and RH the vapour pressure for each level is $e = \mathrm{RH} \cdot e_s(T)$, and from it

$$r = \epsilon \cdot \frac{e}{p - e}, \quad \epsilon \approx 0.622$$

$\ln r$ is interpolated, because the mixing ratio decreases approximately exponentially with height; at the target elevation, vapour pressure and relative humidity follow from it again. The reason for this detour is the exponential temperature dependence of the saturation vapour pressure: a linearly interpolated relative humidity produces profiles at layer and cloud boundaries that do not exist, whereas the mixing ratio measures water vapour against the dry air mass and is conserved under lifting.

Outside the available levels no value is produced. If a target elevation lies below the lowest or above the highest level that the model provides, no statement is made for that elevation.

3.6.3 Pressure Levels Below the Model Surface

A pressure level counts only if it lies at least 100 m above the model's orography. Below that it is not a computed air column but the result of the pressure-level interpolation, which is extended into the model topography so that maps over mountains show no gaps.

The difference is measurable. At the Zugspitze, 950 and 925 hPa lie below ground in every model; there the relative humidity on both levels is identical in 253 out of 253 hours and the temperature difference is never smaller than +1.20 K. Over Munich, where the same levels are free atmosphere, equal humidity occurs in 6 out of 240 hours and the difference reaches −6.20 K — inversions appear there, below ground they do not.

This is precisely the reason for the limit: an extension downwards from the grid surface knows neither cold-air pool nor inversion. It would be too warm in the valley and show rain where snow is falling.

The 100 m are a technical quality threshold, not a meteorological quantity, and are not displayed anywhere. They keep a distance from the layer in which the level nominally lies above the orography but is already shaped by the model's near-surface adjustment. With 50 m, the GFS 850-hPa level at the Zugspitze — orography 1455 m, level height per hour 1492 to 1553 m — lay exactly on the threshold and switched sides from hour to hour.

Which level remains as the lowest therefore depends on the site and not only on the model. The orography is determined for each model in **two steps**.

The elevation parameter does not only adjust the surface variables, it also shifts the **grid cell**: the service searches the surroundings for the cell whose terrain comes closest to the requested elevation. At the Zugspitze, the same coordinate without an elevation returns the ICON EU cell 47.4375/11.0 at 1638 m, with the summit elevation, by contrast, 47.375/11.0 at 1804 m — and the pressure levels come from the second. As measured, they match there in 48 out of 48 hours and in the first in not a single one.

The response states under elevation only the requested value, not the terrain. Therefore the site elevation is first used to ask **which** cell applies, and then this cell is queried without an elevation for its terrain. This is done for each model individually, because a response for several models carries only one coordinate.

Without this detour, the orography at four of six Alpine summits would come out 111 to 524 m too low (Table 28), namely for ICON EU, ICON D2 and AROME — the three models whose cell moves at all. On the coarse grids it stays put. Too low is the unsafe direction: the lower limit of the profile would lie too far down, and a level below the model terrain would get through.

How far the orography deviates from the actual site is set out in Section 16.10: at a summit it always lies below, in a valley always above. Both follow from the same smoothing.

4 Multi-Model Blend and Weighting

The six forecasts are deterministic runs, not members of a probabilistic ensemble generated by perturbing initial conditions or physics. Their weighted combination is referred to as the **multi-model blend**; it is not a calibrated probabilistic ensemble, and the inter-model range is not an uncertainty interval. The word ensemble is reserved for the probabilistic systems behind the precipitation probability (Section 12.3).

4.1 Base Weights

The weighting follows the resolution over the target region.

Model	Weight	Rationale
ICON D2	2.0	Highest resolution over the Alpine region
ICON EU	2.0	High-resolution regional model
AROME	2.0	Regional model for the western Alps
ECMWF IFS	1.5	Global reference model
AIFS	1.5	Medium range
GFS	1.0	Low resolution over the target region

Table 7: Base weights w_m. Sum with all models active: 10.0.

The values are prescribed and not derived from verification data (Section 16.5). Resolution serves as a proxy for skill — a common starting point, but not a measured relationship. Without our own verification it remains open whether this weighting performs better in the Alpine region than equal weighting of all models. It is thus a testable hypothesis, not demonstrated model skill.

These weights therefore apply **only to site estimates** — that is, to variables for which the question is which value is correct at this point. They do not apply to consensus questions; which these are is stated in the following section.

4.2 Lead-Time-Dependent Reweighting

From a lead time of three days onwards, $w_{\text{GFS}} = 0.5$ and $w_{\text{ECMWF}} = w_{\text{AIFS}} = 2.0$; ICON EU keeps 2.0.

The direction of this adjustment is supported by the DIMOSIC comparison (Magnusson et al.[4]): there the IFS shows the smallest error for most variables and regions, with ICON in second place. In addition, the high-resolution regional models are mostly out of range from day 3 onwards (Table 2).

The support is, however, weaker than the figures suggest. DIMOSIC compares seven **global** models; ICON D2 and AROME are not in the comparison, and their weights cannot be justified from it. Moreover, the core finding of the study is that the initial conditions contribute more to differences in skill than the model formulation (Section 2.4). The reweighting is therefore an extrapolation from a related question, not a derivation.

The reference is the lead time from the first forecast day, not the index in the data array, which also contains past days. All views (10 days, 48 h, 3 h) use the same conversion.

4.3 Aggregation Rules

The method combines models in three ways, and the choice depends on what the respective variable expresses.

Site estimates — freezing level, wind, condensation level, precipitation, snow, rain, wet-bulb temperature, sunshine duration, snow depth and boundary-layer height — use the weighted mean

$$\overline{x} = \frac{\sum_m w_m \cdot x_m}{\sum_m w_m}$$

over all models m with a valid value. Here the question is which value is correct at this point, and for that resolution is a defensible proxy.

Consensus questions, by contrast, count models instead of weighing them: the vote on the weather code, the selection of the most severe state in the block together with its duration, the phase consensus, the precipitation probability, and the indicators and thresholds of the thunderstorm score. They do not ask for a value but for how many models see the same thing. Measured with weights, grid spacing rather than agreement would decide there: two high-resolution Alpine models would add up to 4.0, ECMWF and GFS together to 2.5 — with the same number of agreeing models. The thresholds of the cascade are therefore expressed in numbers of models and not in sums of weights; they are uniformly set at two models — the outlier filter of the vote, the fallback from rank 80, the phase consensus, the snow symbol and the thunderstorm adjustment.

One variable in the consensus domain deliberately remains weighted: the **coverage** of precipitation (Section 11.3). It is not a threshold but a fraction — the sum of weights of the models with precipitation divided by the sum of weights of those with any value at all. Where a threshold asks whether enough models agree, a fraction asks how much of the available opinion is wet; there a model's skill in the Alpine region is the right currency. The justification is given at that point.

Safety-relevant extremes are carried unweighted as extreme values: the maximum and minimum temperatures are the absolute maximum and minimum over the active models, as are the apparent temperature (minimum) and CAPE (maximum). The gust follows a rule of its own that combines model values with an elevation-dependent lower bound (Section 8.4).

The temperature is displayed as a curve of the site estimate. Each hour carries the weighted mean over the active models, drawn in two colours — warm above the freezing point, cold below it — with a fill towards the freezing point. The highest value of each day (06–18 h) and the lowest of each night (18–06 h) stand as numbers; a night labelled per calendar day appeared twice, once before and once after midnight. Two numbers closer than a number is wide never both remain. The scale spans the values shown, at least 4 °C. The extremes of the models remain in the data but are not drawn: a frosty night forecast by a single model is no longer visible in the display. Whether to show them as a band was examined and rejected — without explanation the band read as noise around the line.

4.4 Model Families

Six model runs are not six independent opinions. ICON D2 obtains its ensemble boundary conditions from the parent ICON system; at short range the two are a time-staggered approach and not two voices.

Where the **breadth** of agreement matters, one source per family therefore counts: ICON D2 as long as it supplies data for the variable in question, thereafter ICON EU. The selection thus follows the horizon of the current model run, not a fixed number of hours — that would shift with the run's age. This affects the condensation level and the displayed data basis: the user sees five forecasting approaches, not six models.

ECMWF IFS and AIFS remain separate. They share the initial analysis but then follow the atmosphere by fundamentally different approaches — physical integration versus a data-driven model. Their errors are correlated but not redundant; halving them would be a prescribed correction weight without foundation.

When **detecting** a hazardous state, the family rule deliberately does not apply. That ICON D2 and ICON EU report the same thunderstorm is an indication, even if both come from one branch: they represent different resolutions and ranges. The cascade checks the number of supporting model runs there, not their statistical independence. The asymmetry is intended — conservative in detection, honest in stating breadth. Blocks in which a severe state comes exclusively from one origin are

recorded in the data model, so that it can later be checked whether they are more often false alarms or more often early warnings.

The assignment follows the model architecture, not a measured error correlation. An evaluation at the Zugspitze over 312 hours supports it for temperature — ICON D2 and ICON EU differ there by 1.34 K on average, all other pairs by 1.70 to 2.87 K. For precipitation the same evaluation shows no difference: all pairs lie between 0.21 and 0.28 mm. The rule thus prevents structural double counting, but it is not a weighting calibrated from error correlations.

5 Freezing Level

5.1 Method

The freezing level is not taken from the field `freezing_level_height`. That field is not defined consistently across the models: measured at two sites over 312 winter hours each, ICON D2 never falls below its grid elevation and lies exactly on it in 299 and in all 312 hours respectively — there the field is truncated at the model topography. ICON EU falls below the same limit in 280 of 312 and in all 312 hours, GFS reports zero metres in 104 and 73 hours respectively. An average over these fields would combine three different quantities.

Instead, the height is sought at which the temperature in the model's profile (Section 3.6) changes sign. Between two adjacent levels with $T_a > 0$ and $T_b \leq 0$:

$$\mathrm{FZ} = H_a + \frac{T_a}{T_a - T_b} \cdot (H_b - H_a)$$

The search proceeds from bottom to top, and the **first** crossing found counts. In an inversion the zero line is crossed several times; for a tour the lowest one is decisive, because the cold part of the mountain lies above it.

All six models provide pressure levels and can carry the freezing level. Whether one does so in a given hour is decided by geometry: the zero line must lie between two levels that lie above the model's own terrain. Measured over 360 winter hours at fourteen Alpine sites, AROME contributes no freezing level at ten of them, because its column stays below zero throughout; in summer it contributes throughout at ten sites.

5.2 Validity Conditions

No extrapolation takes place. If the freezing level lies below the lowest or above the highest level that the model provides, that model contributes no value for that hour. The lower limit is thus model-dependent and follows the level ladder in Table 6, not a prescribed number.

No value is produced either if two levels are less than 50 m apart in height or if a model provides fewer than two complete levels — complete meaning temperature, humidity and geopotential height together. Half a level supports nothing.

The block value is the weighted mean over the hours with a valid sign change, the blend value the weighted mean over the models that have one.

5.3 Interpretation of the Displayed Row

Without extrapolation there are situations in which no model supplies a height. In winter this is the rule: the freezing level then lies below the lowest level above ground, at the Zugspitze it was found in the profile in not a single one of 312 hours. The profile then knows only that there is frost at every elevation band, and a row of dashes and frost symbols explained nothing to the reader.

The displayed row therefore always carries a number. Two sources stand in order:

Profile At least 60 % of the weight of the models covering this time finds a sign change in their own profiles; the value is the weighted mean over them. It describes the stratification of the day and is used whenever it exists — in summer, high above the site, where it matters for high-alpine tours.

Estimate from the site Only where the profile finds no line, it is carried from the temperature at the site — for an hour the weighted mean that the temperature curve shows, for a six-hour block the weighted mean of each model's midpoint between its lowest and highest value — with the lapse rate of the ICAO standard atmosphere, 0.65 K per 100 m, the same as for the displayed snowfall level (Section 5.4.2). It never drops below sea level.

A profile that disagrees with the temperature at the site is kept. Above freezing at the site and a line below it, or the reverse, can come from an inversion, from the elevation adjustment or from different sets of models; it does not show that the profile is wrong. The disagreement is recorded with every value, not removed. In the summer recordings it occurs in 5 of 52 blocks at the Zugspitze and 11 of 52 at the Grossglockner, with a site temperature between −2 and +3 °C and the line 6 to 470 m from the summit; most are night blocks in which the summit lies just below freezing while the free atmosphere above is warmer — radiative cooling at the surface, which a single site value cannot express. In the winter recordings the profile carries the line in 6 and 0 of 52 blocks, and the estimate the rest.

For an hour the estimate agrees with the temperature curve by construction; for a block it lies above the site where the block's site value is above freezing and below it where that value is below. A test checks this against all four recordings, loaded with the orography and elevation fields as in the application. The price is twofold: the source changes with the situation, and the number may jump where it switches; and the estimate does not know the stratification of the day. The source is carried with every value.

The data continue to carry the profile states as before: a height, frost at every elevation band when at least 75 % of the model weight report it, or no statement. The frost condition is tested from the same sources as the phase — native profile above the model surface, otherwise the surface fields converted to the elevation band.

The reference for the 60 % is the models that cover the respective time, not all six. A model whose run ends is not missing — it simply does not reach that far. ICON D2 ends after about two days, ICON EU after about five; thereafter the medium-range models carry the row alone, and computed against the full weight the second forecast week would lose the number entirely. As measured, this is 53 versus 91 % of the hours, although ECMWF and AIFS find a freezing level there in every hour of their run.

Two conditions decide the denominator, and they ask different things. The site temperature establishes **temporal** coverage: the run extends into this block. The lowest usable level establishes **vertical** coverage: the column reaches down to the site. Only a model that satisfies both counts — either with a height or as a dissenting vote.

A missing sign change is a statement about the site only if the column reaches it. If the column begins above, the layer in between remains unobserved, and that is exactly where the freezing level may lie; "I see no freezing level" would then only mean "I am not looking there". Such models appear neither in the numerator nor in the denominator. This affects an elevation band between about 3100 and 4400 m: high enough for the cell chosen for the requested elevation to push the 700-hPa level below ground, and low enough for the next level to lie above. At the Eiger (3967 m) the column of ICON D2 and AROME begins at about 4375 m and thus 400 m above the summit; on average 2.3 of 10.0 weight units are excluded there, at the Grossglockner 1.5, at the Zugspitze and Mont Blanc none.

A technical failure, by contrast, remains in the denominator — otherwise an error would improve coverage. How much weight was excluded for lack of column coverage and how broad the basis is relative to all models are carried along as separate metrics, without deciding the display.

For the block, the majority of hours decides. A tie counts as **not** covering: three out of six hours are not a majority, and the consequence of an affirmative answer would be the more severe one — as a dissenting vote, the model would depress coverage without its column predominantly reaching the site.

Two reference elevations, two questions. Here the condition is $z_{\text{lowest}} \leq z_{\text{site}}$ without margin — the question is whether the column reaches down to the site. For the column maximum of the new-snow amount it is $z \geq z_{\text{orography}} + 100$ m — there the question is whether an individual value was computed above ground at all (Section 3.6.3). A level can be atmospherically sound and still begin 400 m above a summit; and it can lie above a valley site and still be below that site's model terrain. The two rules therefore measure against different surfaces, and the 100 m are a masking buffer, not a site tolerance.

5.4 Rain–Snow Line

Alongside the freezing level, a second height is carried: the height below which precipitation falls as rain. Where the pressure-level profile supports it, it is derived from that profile with the same wet-bulb method as the phase per elevation band (Section 13.1). Where the profile does not support it, lower-ranked sources take its place (Section 5.4.1). These are proxies, not equivalent diagnostics: they use a different threshold or a coarser vertical basis, and the source is retained with every value.

The sign change is sought not at $T_w = 0$ °C but at the midpoint of the transition zone of the phase curve (Section 6.2):

$$T_w^{\text{mid}} = \frac{T_w^{\text{snow}} + T_w^{\text{rain}}}{2} = 0.75 \text{ °C}$$

There, by definition, half of the precipitation is solid — the usual definition of the snowfall level in mountain forecasts. The value is **derived** from the two zone limits and not set alongside them: if they shift, the line moves with them.

The line is therefore not the height above which only snow falls — that would be $T_w = 0$ °C, measured to lie about 120 m higher. It is the middle of a transition zone that is about 240 m thick in the Alpine region. Within the zone, the precipitation bar at the site is two-coloured.

It lies below the freezing level, and the distance is not a constant: over 156 blocks at three Alpine summits, 421 to 587 m on average. Both heights are therefore carried separately. The numeric row shows the freezing level of the air temperature — the quantity that avalanche bulletins and mountain

weather forecasts state and against which experience is calibrated. The displayed snowfall level is not this line but an estimate from the site (Section 5.4.2).

5.4.1 Source of the Elevation Estimate

The profile supports the line only as long as it lies above the lowest usable level. In winter it lies below: the transition layer then sits between the model surface and 700 hPa, and no model provides a level there. At the Zugspitze, the profile found the line in January over 504 hours in 0 % of the hours (ICON EU, ICON D2, GFS) and in 7 % (ECMWF); in August the figure was 65 to 100 %. This is not due to the margin — with 0 m instead of 100 m it remains at 0 to 14 %.

A line that exists in summer but not in the ski-touring season is no line. For each model and hour the following order of precedence therefore applies, and the block value is formed from **one** height per model:

1. **Our own pressure-level profile**, if it finds a sign change above the orography. Transparently computed and anchored to native levels.
2. **`snowfall_height`**, the snowfall level that ICON EU and ICON D2 diagnose themselves. Available in 504 out of 504 hours in the same January period. It addresses a related question — where falling precipitation turns to snow — rather than the dry freezing level, but at a threshold of 1.3 °C instead of 0.75 °C. ECMWF, GFS and AIFS do not provide the field.
3. **The surface fields queried per elevation band.** The fields of each elevation band yield a wet-bulb temperature; the lowest sign change between two adjacent bands is sought. This is not a downward extension from the summit value but an interpolation between independently queried diagnoses. There is no continuation beyond the lowest and highest band.
4. Otherwise no line.

The source is carried along in the data for each model. Without it, a jump at the end of the ICON range could not be distinguished from a change in the weather.

The lower-ranked sources are coarser than the profile, and that is the reason for their order. `snowfall_height` is a provider diagnosis whose derivation is not fully documented. The band fields rest on points 1000 m apart; if the service's downscaling applies the same lapse rate to all elevations, the sign change found is more a systematic shift than a represented stratification, and it does not capture valley inversions. The displayed height is an orientation value, not a point-accurate measurement.

Frequent availability is no mark of quality. The order of precedence takes effect most strongly precisely where our own profile carries least — and these are the same models and the same sites at which the column does not reach the location. At the Eiger (3967 m) the values of ICON D2 come from `snowfall_height` in 138 hours and from the profile in not a single one; for ICON EU, 139 hours from the provider diagnosis stand against 68 from the profile, while ECMWF, AIFS and GFS carry the profile throughout. AROME supplies no height there by any of the three routes.

Two things follow from this. A value from `snowfall_height` belongs to the model's grid point — for ICON D2 at the Eiger at 3479 m — and not to the summit; it is therefore not an equivalent vertical diagnosis of our own. And any assessment of the quality of this line must be separated by source, otherwise it averages three methods into a number that applies to none of them.

The switch between the first two sources carries a systematic offset. In the summer recordings, where both are available in the same hour, the profile line lies above `snowfall_height` by a median of 84 m (ICON EU, Zugspitze, 201 hours; 10th–90th percentile 6 to 176 m), 94 m

(ICON D2, Zugspitze, 71 hours) and 53 m (ICON EU, Grossglockner, 164 hours). Evaluated at 1.3 °C instead of 0.75 °C, the same profiles reproduce `snowfall_height` to within a median of −29 to +14 m; the offset is therefore the difference in threshold, not a difference in the atmosphere. In winter, when only `snowfall_height` is available, the line thus lies roughly 50 to 100 m lower than the profile definition would place it. No correction is applied: a fixed shift derived from two summer sites would be a hidden calibration. Where the source changes within a time series, the jump must not be read as a change in the weather.

5.4.2 Displayed Snowfall Level

The snowfall level stands in a row of its own, in metres rounded to 100 m, ahead of the temperatures. It must not contradict the precipitation bar and the amounts beside it, and the line of this section cannot guarantee that: in winter it rests almost entirely on `snowfall_height` — at the Zugspitze and Mont Blanc 622 of 669 model-hours with a line — that is, on two models and a provider diagnosis for the grid point, while the phase at the site comes from our own wet-bulb temperature over all six. At a site close to the line the two disagree: at a 2115 m summit the line stood at 2400 m while 8.3 cm of snow fell at the site.

The displayed value therefore starts from the same quantity as the colour of the bars. A wet-bulb temperature at the site is carried upwards or downwards with a fixed lapse rate of 0.65 K per 100 m, the temperature gradient of the ICAO standard atmosphere, to the midpoint of the transition zone:

$$z_{\text{snow}} = z_{\text{site}} + \frac{T_w - T_w^{\text{mid}}}{\Gamma}, \quad \Gamma = 0.0065 \text{ K/m}$$

Where precipitation falls, T_w follows its solid fraction at the site. Within the transition zone the fraction itself gives the wet-bulb temperature; entirely solid or entirely liquid precipitation bounds it to the edge of the zone, and the block mean carries the distance beyond. The block mean alone would not do: it includes the dry hours, and rain often falls in the warmer ones. Without precipitation the block mean decides. The value never drops below sea level.

By construction the row thus lies below the site where it snows there, above it where it rains, and close to it where both fall. A test checks this against all four recordings. The price is the assumption: the fixed lapse rate does not know the stratification of the day, and in an inversion the estimate is wrong. It is a prescribed parameter (Section 15), not a measured one.

5.5 Snowfall Level in ICON

With `snowfall_height`, ICON EU and ICON D2 supply a diagnosis of their own: the height at which the wet-bulb temperature exceeds 1.3 °C on descent. The field is also defined outside precipitation hours, because it follows from the humidity structure of the column and does not depend on falling precipitation.

Snowfall level and freezing level are **different quantities**, not variants of the same one. One answers where falling precipitation turns into snow, the other where the air reaches freezing point. They are therefore carried separately: the displayed freezing level is based on the interpolation in Section 5.1 for all models with pressure-level data. `snowfall_height` enters only on the phase side — there the wet-bulb temperature is the more appropriate quantity, because it is closer to the temperature of a falling hydrometeor. On the phase side, the field carries the rain–snow line when the pressure-level profile does not find it (Section 5.4.1).

The justification comes from the literature on phase determination: the wet-bulb temperature is closer to the surface temperature of a falling hydrometeor than the air temperature and takes humidity into account. Jennings et al.[13] use it as the central variable of their study; Fontrodona-Bach et al.[12] explicitly recommend it for climates with frequent precipitation near freezing point and name the Alps in this context — an argument for the marginal range, not evidence of general superiority (Section 2.2).

Three heights must be strictly distinguished: the freezing level as the height with $T = 0$ °C, the wet-bulb freezing level as the height with $T_w = 0$ °C, and the operational snowfall level as the height above which snow is expected at the target level. In saturated air the first two almost coincide; in unsaturated air the wet-bulb freezing level lies several hundred metres lower. Depending on the definition, the third can lie above or below the second — some operational systems place it not at $T_w = 0$ °C but at a positive threshold.

The difference between these heights is therefore not constant. It depends on the temperature and humidity profile, pressure, precipitation intensity, melting and evaporation processes, and on the definition of the respective diagnosis. Several hundred metres are physically possible in unsaturated air. The quantities are therefore not converted into one another by a fixed height offset.

5.6 Wet-Bulb Temperature at the Site

The ICON diagnosis covers two of six models. For the others, the wet-bulb temperature is computed by us — hourly, for the site elevation, from temperature, relative humidity and surface pressure.

The reason lies in a peculiarity of the data preparation. The service adjusts temperature, dew point, humidity and pressure to the requested elevation; it does **not** adjust the partitioning of precipitation into a liquid and a solid fraction. This partitioning describes the grid elevation of the respective model, and at high summits that lies far below the site — at the Zugspitze between 654 and 1619 m lower for the cells the processing uses, depending on the model (Table 28, Section 3.6.3).

The wet-bulb temperature is thus the only phase variable that applies to the site. The psychrometric equation

$$e = e_s(T_w) - \gamma \cdot p \cdot (T - T_w)$$

with the saturation vapour pressure over water, $e_s(T) = 6.112 \cdot \exp(17.62T/(243.12 + T))$ hPa, and the psychrometer coefficient $\gamma = 6.65 \cdot 10^{-4} \cdot (1 + 9.44 \cdot 10^{-4} T_w)$ K^{-1}, is solved numerically, by 40 bisection steps between dew point and air temperature — an interval that is certain to enclose the solution. The formulation over water is used below 0 °C as well; for the phase transition, which lies above −2 °C, the difference to an ice-bulb formulation is small. If a model supplies no surface pressure, the direct formula of Stull[19] applies and is flagged as an approximation. All six models provide surface pressure; the branch is thus a fallback for individual gaps, not a rule for a particular model.

Pressure is the reason why the approximation does not suffice throughout. It was formulated for sea level; at 2962 m the surface pressure is about 710 hPa, and with otherwise identical values the wet-bulb temperature there is up to 0.56 K lower. At the phase transition this shifts the snowfall level by about a hundred metres. At humidities of 40 % and above, the two computations remain within 0.8 K of each other; in the cold, dry corner that Stull explicitly excludes, they diverge by 2.4 K.

5.6.1 Limits of the Approximation

This corner is not a side remark. Stull states the formula for −20 to 50 °C and 5 to 99 % relative humidity; below that it departs from physics. At −25.1 °C and 12 % humidity it yields −22.69 °C — a wet-bulb temperature **above** the air temperature. That cannot be: a falling hydrometeor cools by evaporation, it does not warm. Measured over four fixtures and 6576 hours, this occurs in 655 hours, by up to 4.37 K.

The stated range alone does not catch the overshoot: 40 of these hours lie within it. Both are therefore checked together — the range and the bound $T_w \leq T$ itself. If either fails, **no** value is reported for that hour. In particular, no value capped at the air temperature: that would look like saturated air and would not be.

This is exactly the cap that the data service applies. Its field `wet_bulb_temperature_2m` reports, in 100 % of the hours in which the same formula overshoots, a wet-bulb temperature equal to the air temperature — at Mont Blanc in 34.6 % of all winter hours, at humidities of 7 to 73 %. There the field is no longer a wet-bulb temperature but the visible trace of an overshoot.

Downwards, by contrast, values are not discarded but brought back. In saturated air — 96 to 100 % humidity — the approximation falls below the dew point by up to 0.37 K. That is its ordinary error, whose stated mean is 0.3 K, and there the correct answer is known: at saturation, dew point, wet-bulb and air temperature coincide. Discarding would hit exactly the hours in which precipitation falls. The two humidity ranges do not overlap, and the distinction rests on this.

The phase remains determined even without a value. Because the wet-bulb temperature never lies above the air temperature, precipitation under frost at the site is solid as far as any site variable can tell — a bound, not an estimate. Supercooled rain out of a warm layer aloft lies beyond this bound; no site variable shows it, and only the ECMWF precipitation type can report it (Section 14.2). The failure of the approximation affects only the cold, that is, the situation in which the answer is known anyway; the computed new-snow amounts of the winter cases do not change as a result.

5.6.2 Why the Provider Field Is Not a Reference

It is tempting to measure our own computation against the service's field. That does not hold, for a fundamental reason: the service computes Stull without pressure, our production value solves the psychrometric equation with pressure. The distance between the two measures the pressure effect, not correctness.

Case	Median	p95	Maximum	above 0.3 K
Zugspitze winter	0.343 K	0.586 K	1.916 K	60.2 %
Mont Blanc winter	0.300 K	0.875 K	3.092 K	49.3 %
Zugspitze August	0.195 K	0.348 K	0.933 K	11.0 %
Grossglockner August	0.200 K	0.319 K	0.722 K	7.6 %

Table 8: Distance of the psychrometric computation from the provider field. An alarm at 0.3 K would fire in every second hour in winter — it measures elevation, not error.

As a regression check against our **own** approximation, by contrast, the field is useful: same formula against same formula, measured at 0.03 K median and 0.15 K maximum, almost identical across all four cases. This check triggers if the implementation, the field mapping or the service's method changes. The production value is given no threshold against the field.

The entire phase decision is thus based on the wet-bulb temperature: from it the solid fraction of precipitation is derived (Section 6.2), and from that the displayed snow amount (Section 7). It is the only variable in the chain that applies to the site elevation and at the same time decides the phase.

6 Precipitation Phase

6.1 A Site-Specific Classification

The phase is determined at exactly one point: in the partitioning of precipitation for the site elevation, from the wet-bulb temperature. The following section describes it. Everything else — which symbol a block carries, whether an amount counts as snow or as rain — reads this result instead of deciding it anew.

Condition	Phase
no precipitation	–
ECMWF, site at most 300 m above the grid: freezing rain	Rain
ECMWF, site at most 300 m above the grid: wet snow	Mixed
ECMWF, site at most 300 m above the grid: ice pellets	Snow
solid fraction above the threshold, liquid below	Snow
liquid fraction above the threshold, solid below	Rain
both above the threshold	Mixed
both below, remainder without wet-bulb temperature above	Unresolved
both below	Trace

Table 9: Phase decision per model; the first applicable condition determines the result. The threshold is the same as everywhere in the cascade, 0.1 mm.

The ECMWF precipitation type takes precedence because it is not a second derivation of the same quantity but a model statement of its own: freezing rain and wet snow are states that a wet-bulb temperature alone does not express. It applies unless the site lies more than 300 m above the ECMWF grid: the type describes what arrives at the grid's own terrain, and at summits that terrain often lies more than a thousand metres lower — at the Zugspitze 1619 m. Freezing rain in the valley is not freezing rain at the summit. The rule is one-sided: where the grid lies above the site, as it regularly does in valleys, freezing rain reported for the higher terrain holds all the more for the cold air below, and the type stays — for freezing rain only while the site itself is at or below 0 °C, since it freezes on contact with frozen ground. The 300 m are prescribed, roughly the thickness of the snow–rain transition zone. At Alpine summits the ECMWF grid lies 667 to 2732 m lower (Table 28); there the rule switches the type off in effect. Ice pellets fall as a solid and are therefore assigned to the solid phase; what follows from them on the mountain is stated by the type (Section 14.2).

If both fractions lie below the threshold although precipitation is falling, this is not mixed precipitation but one of two other situations: a trace amount, or an amount falling in hours without a wet-bulb temperature. Neither is a statement about the phase, and a model that cannot say anything about the phase should not take part in a vote on it.

The phase decides which symbol a model supports, not whether its amount counts: the displayed snow and rain amounts are weighted means over all models with valid precipitation, including those that report zero (Section 14.4). Trace and unresolved amounts are therefore carried as states of

their own, separate from mixed precipitation: they count towards the amounts, but support neither a snow nor a rain symbol.

6.2 Partitioning at Site Elevation

The models output their own partitioning into rain and snow for the elevation of their grid point, not for the site (Section 16.10). At high summits that grid point lies several hundred to over a thousand metres lower, and rain falls there where snow arrives at the summit. The amount is therefore partitioned anew — from the wet-bulb temperature at the site (Section 5.6), which applies to the requested elevation.

The model's total amount remains unchanged. All that is determined is which fraction of it arrives at the site as solid:

$$P_{\text{solid}} = P_{\text{total}} \cdot f_{\text{snow}}(T_w), \quad P_{\text{liquid}} = P_{\text{total}} \cdot (1 - f_{\text{snow}})$$

f_{snow} blends linearly between two limits: below 0.0 °C wet-bulb temperature precipitation is treated as fully solid, above 1.5 °C as fully liquid. A hard threshold would let a tenth of a degree decide between zero and the full amount. The limits are **prescribed, not calibrated**: they are of the same order as operational diagnoses — 0.5 °C at the US National Weather Service, 1.3 °C in ICON — but have not been tested against any observation and are not fitted to another provider's snowfall level.

The partitioning conserves the amount: the solid fraction, the liquid fraction and the remainder for which no wet-bulb temperature is available together give the total amount. The third fraction is carried separately and is not dropped.

It depends only on the thermodynamics of the respective model and thus **not** on the user's model selection.

This partitioning determines the displayed values. The displayed snow amount is derived from the solid fraction (Section 7), the phase of the block from the ratio of the two fractions (Section 6.1). It is the only path by which a snow amount is produced; its distance from the models' own partitioning is quantified in Section 16.10.

6.3 No Additional Temperature Threshold

The spatial variation of the rain–snow threshold demonstrated by Jennings et al.[11] is not reproduced. In its place stands a fixed transition zone in the wet-bulb temperature at the site (Section 6.2), the same everywhere. A second, independent check does not take place: the wet-bulb temperature at the site is the method's only phase decision (Section 6.1).

A threshold map of our own would be a calibration that has not been validated for the Alpine region. The expected gain is moreover doubly limited: by the upper bound on achievable accuracy shown by Jennings et al.[13], and by the finding of Fontrodona-Bach et al.[12] that shifted thresholds do not consistently improve model skill.

7 Snow-to-Liquid Ratio

7.1 Scope of Application

Every displayed snow amount is produced in the same way. The water basis is the model's total precipitation, of which the solid fraction according to the wet-bulb temperature at the site (Section 6.2) is taken, converted with the snow-to-liquid ratio:

$$\text{Snow [cm]} = \frac{P_{\text{total}}[\text{mm}] \cdot f_{\text{snow}}(T_w) \cdot \text{SLR}(T_{\max}) \cdot f_{\text{wind}}}{10}$$

The input to the SLR curve is the maximum of the air column **above the site** — not the temperature at 2 m alone, but not without it either.

A level must satisfy **both** conditions for this, and they ask different things. It must lie above the site, otherwise it would describe air through which the snow no longer falls — that is a question of reference. And it must lie at least the masking buffer of 100 m above the model surface (Section 3.6.3) — that is a question of validity: below that, its value is extended underground and is not a computed atmosphere. Besides the site temperature, the column thus contains only pressure-level values that are physically above ground.

The second condition takes effect in the **valley**, not at the summit. The cell chosen for the requested elevation regularly lies several hundred metres above the location there, and the 850-hPa level falls between the two heights: above the site and at the same time below ground. In Zermatt (1608 m) this applies to ECMWF in 85 of 168 hours; without the check, the column maximum there would be 3.7 K too warm on average and 9.3 K in the extreme. At five summit sites examined, the case does not occur.

Three variables that might suggest themselves at this point do not carry the computation.

The reported rain is not suitable as a water basis. For ICON D2 the budget does not close there: `rain` plus showers and snow water equivalent lies below the total amount in 83 of 116 precipitation hours, never above it; in the extreme, 7.7 of 11.4 mm are missing. At high summits there is the additional problem that the grid points report snow instead of rain — at Mont Blanc, `rain` sums to 1 mm over an entire winter against 873 mm of precipitation. The estimate therefore fails there.

The snow amount **reported** by the model does not carry the display either. At the data service it is not a modelled depth but the water equivalent times a constant 0.7 — over 115 hours at four summits, without exception 0.700 cm per millimetre. Next to a temperature-dependent curve there would thus stand a flat 7:1, different by a factor of 1.1 to 4.3 for the same weather situation. The reported field therefore serves as a control variable.

Damping at the freezing level does not take place either; the justification is given in Section 7.4.

Even a model without a separate rain field would contribute an amount in this way: the computation requires precipitation, temperature and humidity, nothing more.

The ratio is dimensionless: it states how much new snow results from an amount of water, in the **same** unit. A ratio of 13:1 turns 2 mm of water into 26 mm of snow — that is, 2.6 cm. The precipitation fields are given in millimetres, the display uses centimetres; the division by ten lies in between.

7.2 Temperature Dependence

Crystal habit and deposition density depend primarily on temperature. Libbrecht[20] summarises the physical state of knowledge: between about −12 °C and −18 °C, branched dendrites grow that settle loosely (dendritic growth zone); above and below, plates and columns of higher density form.

The curve used is piecewise linear:

$$\mathrm{SLR}(T) = \begin{cases} 8 & \text{for } T \geq 0\ ^\circ\mathrm{C} \\ 8 + \frac{-T}{2} \cdot 4 & \text{for } -2\ ^\circ\mathrm{C} \leq T < 0\ ^\circ\mathrm{C} \\ 12 + (-2 - T) & \text{for } -16\ ^\circ\mathrm{C} \leq T < -2\ ^\circ\mathrm{C} \\ \text{NWS table breakpoints} & \text{for } T < -16\ ^\circ\mathrm{C} \end{cases}$$

Down to −16 °C this is the Kuchera method: $12 + 2 \cdot (271.16 - T)$ above 271.16 K (−2 °C) and $12 + (271.16 - T)$ below, with T in kelvin. It was developed by E. Kuchera at the US Air Force Weather Agency and is used widely in operational forecasting without a formal publication of its own; the formula is cited here after the description of Rosenow et al.[8], who evaluate it among the operationally used ratios. Below −16 °C the breakpoints of the operational table[7] take over, as set out below.

The input T is the maximum over the 2-m temperature at the site and the pressure levels above it. Which level is included is decided by its geopotential height: it must lie at or above the site elevation, with a tolerance of 50 m for numerical imprecision. The tolerance does not extend the column; it merely prevents a level at site level from falling in or out depending on rounding. If the geopotential height is missing, the level stays out — whether it lies above the site cannot then be decided, and too warm a column is exactly the error that this selection corrects.

The 2-m temperature is thus the starting value, which the pressure levels can only raise. At a high summit under cold air it remains the maximum. At Mont Blanc (4808 m) only the 500-hPa level remains: over 312 winter hours, 850 and 700 hPa lie at a median of 1387 and 2894 m and thus in not a single hour above the site. This is the appropriate consequence of the elevation, not a failure.

Down to −16 °C the curve rises: the colder, the looser the snow. Below that this no longer holds. The dendritic growth zone, in which the loosest crystals form, lies roughly between −10 and −20 °C; at lower temperatures the dendritic structure decreases again and the snow becomes denser rather than looser.

A curve that continues monotonically and is capped at −20 °C does not represent this behaviour. At −16 °C the curve used agrees exactly with the operational table after Cobb and Waldstreicher — both give 26:1 —, and this is the natural transition point, not a prescribed mark. Below it, the curve follows the table's breakpoints down to −24 °C, linearly interpolated, and then runs linearly to the table's 7.2 at −30 °C. The table's two intermediate values at −26 and −28 °C (7.0 and 6.8) are omitted, so that the curve keeps falling monotonically instead of dipping below its end value; the difference is at most 1.3, at −26 °C, and concerns only columns colder than −24 °C throughout:

Column maximum	Monotonic curve	NWS table 2011
−16 °C	26.0	26.0
−18 °C	28.0	23.0
−20 °C	30.0	18.0
−22 °C	30.0	12.0
−24 °C	30.0	8.8

Column maximum	Monotonic curve	NWS table 2011
≤ −30 °C	30.0	7.2

Table 10: The cold branch, set against a monotonically continued curve. Only the temperature shape is adopted, not Cobb's method: that additionally draws on vertical motion and humidity, which are not available here. The table limits an extrapolated overshoot; it is not a curve calibrated for the Alpine region.

7.3 Wind Correction

Roebber et al.[5] list external compaction by wind as one of seven factors of snow density. What is meant is a near-surface process: wind redistributes and breaks freshly fallen crystals at the surface, which increases the density of the uppermost layer.

For the reduction, each model's 10-m wind at the grid point is therefore used. The upper-level wind blended from 10 m, 850 hPa or 700 hPa according to site elevation for display (Section 8.1) does not enter: it describes the flow at summit level, not compaction at the ground. No exposure adjustment applies either, and the factor acts per model before aggregation — only precipitation, temperature and wind of the same model run determine the amount estimated for that model. If the 10-m field is missing, the ratio remains unchanged instead of falling back on another wind variable.

A linear reduction is implemented: no effect below 20 km/h, a reduction of 30 % from 40 km/h, linearly interpolated in between. The thresholds correspond to 5.6 and 11.1 m/s and bracket the order of magnitude of about 9 m/s from which the literature describes wind-induced surface compaction. The linear form and the upper limit of 30 % are prescribed and not calibrated against observations for the Alpine region.

7.4 Why No Damping Is Applied at the Freezing Level

A common addition to the site phase is a second factor: if the freezing level lies above the site elevation, the amount is damped linearly and drops out entirely from an excess of 300 m. The method does not apply it.

Both factors answer the same question — does precipitation arrive here as a solid? Multiplying them would apply the phase fraction twice. More serious is that they contradict each other.

The wet-bulb temperature is always less than or equal to the dry-bulb temperature. The level at which $T_w = 0$ therefore lies **below** the freezing level. A site in between has $T > 0$ and $T_w \leq 0$ — precisely the case for which the wet-bulb temperature exists: snow at air temperatures above zero when the air is dry enough. A damping would suppress there what the wet-bulb temperature correctly identifies.

Applied as a sensitivity test at the fixture sites, over all hours with a solid fraction:

Site	Hours	Amount kept	Amount erased
Zugspitze, January	454	454	0
Mont Blanc, January	519	519	0
Zugspitze, August	195	0	195
Grossglockner, August	596	187	117

Table 11: In winter the freezing level lies below the summit anyway, and a damping would do nothing. In summer at the Zugspitze it would erase every single hour in which the wet-bulb temperature yields solid precipitation.

The 300 m are moreover not a physical value. The freezing level states where $T = 0$ — not how deep or how warm a melting layer is. A layer of 50 m at +0.2 °C and one of 600 m at +3 °C give

the same height and an entirely different effect. What the damping would stand in for — the phase at the site — is computed directly.

The objection remains that a warm layer **above** the site can melt a falling crystal while it is freezing at the site. That is correct, but requires temperature and humidity on several levels together with layer energy — not a height difference. It is noted as a separate project in Section 16.

The freezing level is displayed because it is useful for orientation in elevation; it is not an amount factor. The snow amount therefore does not depend on the user's model selection either.

7.5 Comparison with the Profile-Based Method

Between −16 and −24 °C the curve used follows the breakpoints of the operational table[7] and agrees with it there; below that it differs by at most 1.3 (Section 7.2). Above that it remains the Kuchera line and lies above the table.

T (°C)	topo^alpin	Cobb–Waldstreicher	Difference
−10	20	12	+8
−14	24	22.5	+1.5
−16	26	26	±0
−18	23	23	±0
−20	18	18	±0
−24	8.8	8.8	±0

Table 12: Comparison of the SLR curves. Between −16 and −24 °C the two are identical, because there the curve used adopts the breakpoints of the operational 2011 table[7].

The remaining gap therefore lies in the **warmer** branch, between −16 °C and 0 °C, and there it is the difference between two published methods, not an ingredient of our own.

Only the breakpoints are adopted, however, not the method. Cobb and Waldstreicher[6] determine the ratio from the vertical profile — from vertical velocity and humidity in the ascent region, weighted by the share of layers in the dendritic growth zone. The interface used does not supply these variables per model and time step. The curve therefore remains a temperature heuristic that leans on a profile method at two ends without being one. Assessment in Section 16.3.

8 Wind

8.1 Vertical Assignment and Blending

Site elevation	Source
< 1200 m	10-m wind
1200–1600 m	Blending 10 m → 850 hPa
1600–2300 m	850 hPa
2300–2800 m	Blending 850 → 700 hPa
> 2800 m	700 hPa

Table 13: Elevation-dependent choice of wind source.

In the transition zones, blending is linear with $t = \frac{h - h_{\text{low}}}{h_{\text{high}} - h_{\text{low}}}$: $v = (1-t)v_{\text{lo}} + tv_{\text{hi}}$. The direction is interpolated vectorially via sine and cosine components, since the arithmetic mean of cyclic quantities is unusable.

8.2 Block Aggregation

The speed is averaged arithmetically over the hours of the block. For the direction, the direction at the moment of maximum speed is used: during frontal passages, vector averaging over six hours yields directions that occurred at no point in time.

8.3 Exposure Adjustment

For summits and cols the wind is increased by 40 % and 30 % respectively, to represent the acceleration at exposed points that the model grid does not resolve. Le Toumelin et al.[21] show effects of this order of magnitude for complex terrain.

This refers to the mean wind and the lower bound of the gust derived from it. The adjustment does not act on a model gust: it already contains the turbulent component, and multiplying both would count it twice (Section 8.4).

On the achievable accuracy in the Alpine region there is a measurement campaign by MeteoSwiss: Bugnard et al. [22] compare ground-based remote sensing with model profiles in complex Alpine topography and find a systematic overestimation of wind speed, with a pronounced bias close to the valley floor.

The classification as summit or col comes from OpenStreetMap and does not represent the actual exposure (Section 16.5).

8.4 Gusts

The models supply gusts only for 10 m above their grid point. Above 1200 m, by contrast, the displayed mean wind comes from a pressure level at site elevation (Table 13). The two quantities thus belong to different heights and, because of the grid-point elevations (Section 3.5), often to different places. Four rules follow from this inconsistency.

First, a model gust that lies below the upper-level wind chosen for the same model and the same site is raised to that upper-level wind. As a valley gust it is not a usable statement about the site — in high mountains this is the rule, not the exception. It is not discarded, however: that would make the model drop out of the gust aggregation entirely, although its upper-level wind enters the mean wind, and would systematically shift the gust mean upwards. Raised, every model with a gust field remains represented without its valley gust passing as a statement about the site. A model with no gust field at all does not enter; its missing gust is carried as provenance information and not filled with a substitute value.

Second, the model gusts secured in this way are aggregated as a **weighted mean** over all models with a gust field — with the same model weighting as the mean wind. The maximum over the models is not used: it would place a single-model extreme opinion next to a blended mean and structurally raise the displayed ratio of gust to mean wind.

Third, the exposure adjustment s (Section 8.3) does **not** act on the direct model gust. It corrects the terrain acceleration of the mean flow not resolved by the grid; a model gust already contains the turbulent component. Multiplying both would count it twice. s continues to enter the floor, because the floor is derived from the mean wind.

Fourth, there is a conservative lower bound. With the blended mean wind W_0 before the exposure adjustment s, the elevation-dependent gust factor $g(h)$ and the cap $G_{\max}$ of 1.7:

$$\text{Floor} = W_0 \cdot \min(g(h) \cdot s, G_{\max}), \quad \text{Gust} = \max(\text{model gust mean}, \text{Floor})$$

The floor thus applies not only when no model supplies a gust, but also when the mean of the model gusts remains below it.

The gust factor $g(h)$ is piecewise linear: 1.2 below 1400 m, from there linearly to 1.3 at 2500 m, further linearly to 1.45 at 4000 m and constant above.

Source	Condition	Marking
model mean	above the floor	none
raised	model mean below the floor	none
derived	no model supplies a gust field	tilde

Table 14: Provenance of the displayed gust value. An unchanged model mean counts as modelled even if the grid points lie lower than the site — that is a different statement from "estimated". Also carried along are how many models the mean comes from, how many of them were raised to their upper-level wind, and how many active models supply no gust field at all.

Only the third row is marked. The raised value remains unmarked, because it is the rule of the summit construction and not the exception: at the fixture sites it applies to 97.5 % (Zugspitze) and 32.5 % (Grossglockner) of blocks, while the derived case does not occur there in a **single** block. In addition, a usable model gust is present in it — the marking stands explicitly for its absence.

Unmarked therefore means **anchored in the model**, not **unprocessed**: raising is a noticeable intervention in the model value, but one that starts from a model gust. How often it applies is recorded in the provenance count, not in the symbol.

The marking thus appears only when the model basis falls away: with deselected models or at long lead times. The provenance of the other cases is carried along (Section 14).

Physically based methods presuppose boundary-layer profiles (Brasseur[14]); with the available fields they cannot be implemented. The factor used is empirical (Section 16.5).

9 Lifting Condensation Level

9.1 Calculation

The lifting condensation level is estimated from the dew-point depression at the site with a factor of 125 m per degree. Lawrence[23] gives the approximation

$$z_{\text{LCL}} \approx \left(20 + \frac{t}{5}\right) \cdot (100 - \text{RH})$$

for this (t in °C, z_{LCL} in m), from which a factor of comparable magnitude follows for mountain temperatures. Temperature and dew point are the values the interface provides at site elevation (Section 3.5); the height above ground is therefore added to the site elevation. Since the dew-point depression is never negative, the level never lies below the site.

9.2 What the Variable Does Not Describe

The value states at what height cloud would form if air were lifted **from the site**. It does not state whether the site is in cloud, and it is not a cloud base in the sense of an observed or modelled layer.

This is not a fine point of interpretation but follows from the construction: the first term is never negative, and the site elevation is added. The value can thus never fall below the site. Measured over 4203 model hours at three Alpine summits, it lies on average 324 to 462 m above the summit in blocks with precipitation — and below it in not a single one of the 156 blocks, although the dew-point depression reaches 0.00 K and lies below 3 K in 47 to 73 % of the hours.

A number that is read as a cloud base but cannot be one is more misleading than no number. It is therefore **not displayed.**

ECMWF determines its cloud base from the three-dimensional model profile by searching upwards from the surface for a layer with appreciable cloud fraction and condensate content. Such a variable is not available at the interface; nor could it be substituted from the humidity on pressure levels, because relative humidity is neither cloud fraction nor condensate and the available levels between 700 and 600 hPa are more than 1000 m apart.

9.3 Application

The variable controls one rule of the symbol cascade (Section 11.3): a cloudy symbol is upgraded to "partly cloudy" if the condensation level lies more than 1500 m above the site and no precipitation is falling. The margin corresponds to a dew-point depression above 12 K — air this dry rules out cloud at the site, and the height condition carries the statement on its own.

Besides the mean, the minimum over the hours of the block and the number of hours below the site elevation are also computed per model. Both are averaged over the models with weights — the minimum as well, because an extremum over the models would let a single one decide the block. They serve diagnostics; none of the three variables is displayed.

10 Thunderstorm Risk

10.1 Composition

Indicator	Threshold	Points
CAPE	> 500 J/kg	+2
	> 200 J/kg	+1
Convective inhibition	< 25 J/kg, if CAPE > 200 J/kg	+1
Lifted index	< -4	+2
	< 0	+1
Showers, strongest model	≥ 0.1 mm	+1
Convective cloud top	one run > 7000 m	+1

Table 15: Base points of the score, at most 7. CAPE enters as the highest value of the models — a precautionary upper envelope. Convective inhibition enters as the lowest model value and only if the CAPE row is reached; the lifted index is a single GFS value. Showers enter as the strongest model, the convective cloud top as a vote.

The lifted index is the temperature difference between a lifted air parcel and its environment at 500 hPa; negative values indicate instability. It is supplied exclusively by GFS (Table 3). Convective

inhibition is supplied as a non-negative magnitude by GFS and, for part of the hours, by ICON EU and ICON D2; small values correspond to a weak cap.

Convective inhibition counts only together with CAPE. Where there is no CAPE, the models report an inhibition of zero: there is no energy to be held back. A low inhibition is therefore not in itself an indication of convection. Its point is granted only if the CAPE row of the table is reached, that is, if at least one model exceeds 200 J/kg in the block. This is an activation threshold for the point, not a lower limit for thunderstorms; smaller CAPE can also allow convection with suitable stratification and lifting.

The rule is a limited logical correction. CAPE and inhibition may still come from different models and different hours of the block once that threshold is reached — a combination that no single model state need show. Whether one model forecasts both conditions in the same hour is carried alongside as a diagnostic with three states: met, not met with valid pairs of values, and not testable because no model supplies both values in any hour. The last state is a statement about the data, not about the atmosphere. Requiring this coherence for the point would change the summer warnings considerably: at the fixtures, the blocks with a score of 4 or more would fall from 15 to 10 (Zugspitze) and from 15 to 9 (Grossglockner), in each case out of 40. Whether these were false alarms cannot be told without observations; the decision is left to a comparison with observed lightning.

The index is a precautionary indication of a situation capable of producing thunderstorms, not a lightning probability. It combines a consensus-based humidity condition (Section 10.3) with stability variables, of which CAPE enters as the upper model value. A high score thus shows the upper margin of the modelled instability; it does not mean that any single model represents all conditions at once.

CAPE depends on the lifted parcel. The variable is vertically integrated, but its value stands or falls with which parcel is lifted: one from the model surface, one from a mixed layer, or the most unstable parcel of the profile. For the input values used, it is not documented per model which of these definitions applies and from which height the initial parcel originates — at the Zugspitze and the Grossglockner, the model surface of the models with a CAPE field lies between 566 and 1619 m below the summit (Table 28). The different model values are therefore not readily interchangeable.

The value is therefore used as an indication of modelled convective instability in the surroundings, not as a site-accurate statement about triggering at a particular summit.

How strongly the score depends on CAPE is shown by a counter-test at the fixtures: if CAPE is set to zero with otherwise unchanged inputs, the number of blocks with a score of 4 or more falls from 15 to 1 (Zugspitze) and from 15 to 0 (Grossglockner), in each case out of 40. CAPE is thus the strongest single contribution; since the inhibition point depends on it, it also carries that point.

The finding measures the internal dependence of the score on this field, not its predictive skill. It rests on two sites and one time window and cannot be generalised either spatially or temporally.

Showers enter as the strongest model: a single model with at least 0.1 mm of convective precipitation grants the point. The threshold is tested per model, not against a mean, which dry models would dilute and which would make the required amount depend on the number of models. The number of models reporting showers is carried as a diagnostic. Requiring two of them, as for the weather-code adjustment, would lower the score in stable precipitation where a single model reports small amounts, but would also withdraw summer warnings at high CAPE; this, too, is left to a comparison with observed lightning.

The cloud top counts votes instead of averaging a height. It comes exclusively from ICON EU and ICON D2, and without convection both report not a missing value but a height of zero. A mean of a

cloud top at 12 km and a zero is not a multi-model statement about a height — it halves it. At the fixtures this cost the point in three of forty blocks each, invariably in the direction of an all-clear.

The indicator therefore asks what it means: does at least one run see deep convection? This corresponds to the treatment of the other score components as consensus questions (Section 4.3). Requiring two votes would, with two possible contributors, be a unanimity rule and would silence the point from the end of the ICON-D2 range. A point from a single vote is therefore an indication and not a consensus; the number of contributing runs, the highest valid cloud top and its spread are carried along as diagnostics.

Missing values take two forms. Without convection ICON D2 reports a height of zero, ICON EU predominantly −500 and otherwise also zero; neither is documented. Only what is positive as a height above sea level is therefore adopted. What the value means remains open: "no convection diagnosed" and "variable not computed" cannot be distinguished at the interface, and both are treated alike — as missing, not as a statement about the atmosphere.

Because the score combines variables with different model bases, the processing also carries the highest score that a **single** model achieves from its own values. Its difference from the displayed value describes how strongly the latter is assembled from partial values of different models. It is carried as a diagnostic variable and downgrades nothing: an automatic downgrade would be a hidden calibration without comparison against lightning data.

In addition, there is an adjustment of two points if the models themselves forecast a thunderstorm: if at least two models report a weather code of 95 or above at any hour of the block — the same threshold as in the symbol vote (Section 11.2) — the adjustment is granted. It rests on the model evidence, not on the displayed symbol, so that changes to the symbol cascade do not shift the score. The result is capped at 8; 7 base points and 2 adjustment points can therefore not produce a higher value.

10.2 Adjusted CAPE Thresholds

The thresholds lie at 200 and 500 J/kg instead of the 400 and 1000 J/kg common in lowland settings, since Alpine convection is triggered at lower CAPE (shallower stratification, orographic and mechanical triggering).

The adjustment is prescribed and not validated for the Alpine region. It points in the same direction as a documented finding: Taszarek et al.[24] compared 3.7 million radiosonde ascents from Europe and North America with the ERA5 and MERRA-2 reanalyses and demonstrated a systematic underestimation of CAPE, particularly in the upper range of values. A comparison of derived CAPE with thresholds from observational climatologies therefore concerns differently scaled quantities.

The finding is transferable only to a limited extent: it applies to reanalyses, not to the operational model output used here, whose CAPE is moreover derived by the data provider (Section 3.3). Whether the same bias occurs there has not been tested. The uncertainty of the thresholds is thus larger rather than smaller than the finding alone suggests.

10.3 Humidity Condition

The score takes effect only if the precipitation probability exceeds 20 %, the shower amount of one model reaches 0.1 mm or the relative humidity exceeds 70 %. Otherwise it is zero. Without this condition, foehn situations with high CAPE and low humidity systematically produce false alarms.

The condition tests unweighted means over the models — like the other score components it asks how many models see a situation, not which value is most plausible at a summit (Section 4.3) —,

whereas CAPE enters as the upper model value. This unequal treatment is intended: for the question whether the atmosphere **can** be unstable enough, a high single value suffices as a precautionary indication; for the question whether the air is moist enough, consensus is required. It does, however, mean that condition and points can come from different models. The diagnostic variable mentioned above is intended for this as well.

10.4 Cloud Layers

Low and high cloud cover are averaged exclusively from ICON EU and ICON D2, since ECMWF and GFS define these variables on isobaric pressure levels rather than on sigma coordinates and the layer boundaries are not comparable. Mid-level cloud cover is defined uniformly at 800–400 hPa and is formed over all active models.

10.5 Achievable Score

A score of 5 does not mean the same everywhere. The lifted index comes only from GFS, convective inhibition mainly from GFS, the convective cloud top only from ICON; the latter ends after about five days (Table 2). From day 6 onwards only 6 instead of 7 base points can therefore be achieved, and the same weather situation yields a structurally lower value in the second half of the week.

The processing therefore carries, alongside the score, the upper limit achievable with the respective data — deliberately without the adjustment, because owing to the cap of 8 it would always raise the upper limit to 8 and thereby conceal precisely the finding that the metric is meant to show. The display shows the score as a violet bar, its height the score relative to the cap of 8; below the elevated level no bar is drawn. The upper limit appears in the tooltip, which presents it as “thunderstorm risk x of y achievable points”.

One bar per column, not per hour. The score combines components that may come from different hours of the block — CAPE from one, showers from another. Computed per hour, a component is missing: at the Zugspitze a block with a score of 3 had a score of 0 in every one of its hours; in the summer recordings, the highest hour fell below its block in 9 and 3 of 52 blocks and never rose above it. A warning that disappears weighs more than a coarser time. In the ten-day view the bar therefore stands per six-hour block. In the 48-hour view each one- or three-hour column shows the score of the six-hour block it belongs to, so that both views warn for the same hours: computed per column, the summer recordings placed an hour in a block with an elevated score in 35 and 41 of 72 cases, but gave the hour itself an elevated score in only 23 and 26. The shrinking data basis is additionally noted in the day header of the ten-day view, which shows the number of models with data against the number of active ones.

10.6 Status of the Score

The score is an uncalibrated indication of thunderstorm potential — not a confirmed risk level and not a probability. It has not been compared with lightning observations. A low value is not shown to be the better forecast, nor is a warning that the score retains shown to be justified.

Winter illustrates the limit. In the two winter recordings the score responds on a single day, a frontal passage with heavy snow: 2 and 3 at the Zugspitze, 2 to 5 at Mont Blanc. There it is built from one ICON vote for a cloud top above 7000 m, showers in one or two models and components from different models; the strongest single model stays one point below the displayed score in each case. Thunderstorms with snowfall occur, and the season alone does not make such a value implausible; the composition shows only that it rests on little.

A rule that requires two models for the shower and cloud-top points is not adopted. For the cloud top only ICON EU and ICON D2 supply data: two votes would mean agreement within one model family, and the condition would become unattainable once ICON D2 leaves its range — a change of the data available over lead time, not merely a stricter test of agreement. The validation planned instead keeps the score unchanged and compares it with lightning data, evaluates in parallel the highest score that a single model reaches coherently within one hour, tests the variants “two shower models” and “two cloud-top votes” separately, and separates hits, false alarms and missed events for winter fronts and summer convection.

11 Weather Symbol

11.1 Severity Ranking

The basis is the WMO weather code according to code table 4677[17].

Rank	Codes	Meaning
90	≥ 95	Thunderstorm
80	85–86	Snow showers
80	66–67	Freezing rain
70	71–79	Snowfall
60	56–57	Freezing drizzle
50	80–82	Rain showers
47	65	Heavy rain
45	45, 48	Fog
40	61, 63	Rain (slight, moderate)
30	51–55	Drizzle
10	1–3	Cloudy
0	0	Clear

Table 16: Severity ranking over the codes according to WMO table 4677[17]. The scale runs in steps of ten so that fog receives a rank of its own between rain and rain showers; it is used exclusively for comparison and never displayed as a number. On the classification of fog and heavy rain see Section 16.7.

11.2 Voting

The code is read for each model at the middle of the block; all six supply one (Table 3). The highest rank determines the result. Within the winning rank, the highest code decides: heavy snowfall (75) beats slight snowfall (71). Without this rule the result would depend on the order in which the models are processed, and that in turn on the order in which the user enabled them. The one exception is rank 80: freezing rain (66/67) comes before snow showers (85–86) there, because glaze ice is the greater hazard; the two still count together for the two-model condition below.

11.2.1 The Block Overrides Its Midpoint

The middle of the block is one hour out of six. At the fixture sites, it carries a milder state than occurs within the block in 33 to 38 % of blocks.

For each model, the most severe state of the block is therefore also carried — by rank, not by code number — together with the duration for which it persists. Across the models it is voted on by the same rules as the midpoint, including the two-model threshold for rank 80.

It may take over only if it persists long enough:

Rank of the state	Required duration
below 60 — drizzle, rain, showers, fog	2 of 6 hours
from 60 — freezing drizzle, snow, freezing rain, snow showers, thunderstorm	1 hour

Table 17: Duration from which a more severe state determines the symbol.

Without duration as a condition, a single thunderstorm hour would displace five calm ones. From rank 60 one hour suffices, because there not displaying it is the greater evil.

The duration counts **only over the models that support the selected state**. An average over all would mix different things: if two models each report one hour of thunderstorm and two others six hours of drizzle, the arithmetic would give almost three hours — for a state that none of the models sees for that long.

Within a rank the rule does not apply: one hour of heavy snowfall does not move the block from slight to heavy snowfall.

From rank 80 — snow showers, freezing rain, thunderstorm — at least two models must agree; otherwise there is a recursive fallback to the next lower rank, for which the same test applies again. Thus no single model can force through a severe report — not even a high-resolution one whose weight alone would reach the threshold. If no model supplies a code, the ECMWF reference code is used, provided ECMWF is selected at all; otherwise the cell remains without a symbol.

11.3 Correction Cascade

Condition	Correction
Snow < 0.5 cm or supported by < 2 models	Downgrade to slight rain or cloudy
Rain showers < 1.0 mm, weight fraction ≤ 50 %	Downgrade to cloudy
Drizzle/rain/showers < 0.5 mm, weight fraction < 40 %	Downgrade to cloudy, otherwise clear
Cloudy, < 0.1 mm, condensation level $>$ POI $+ 1500$ m	Upgrade to partly cloudy
Fair to cloudy, < 0.1 mm, sunshine fraction > 50 %	Upgrade to mostly sunny
Clear to cloudy, block total ≥ 1.0 mm or > 0.5 cm snow	Upgrade to rain or snow showers
Precipitation symbol that contradicts the phase at the site	Mapping to the other phase; towards snow below 0.5 cm to cloudy

Table 18: Correction cascade in actual order of application; each correction works on the result of the preceding one. The amount thresholds of the downgrade rules — 0.5 and 1.0 mm — scale with block length; the thresholds 0.1 mm and 0.5 cm apply unchanged for every block length.

11.3.1 The Symbol Follows the Phase at the Site

The last row of the cascade resolves a contradiction that the preceding ones cannot see. The weather code comes from the model fields and describes the elevation of the respective grid point; the amounts result from the phase decision for the site elevation (Section 6). At high summits the two

diverge: at the Grossglockner 25 % of blocks show snow and no rain — and would carry a drizzle or rain-shower symbol without this rule.

These are not competing statements about the same place, but statements about different elevations (Section 16.10). The rule therefore does not decide which source is right; it ensures that symbol and numbers say the same thing. For a tour decision, "drizzle" instead of "snowfall" is not the same information, and the contradiction cannot be resolved by the reader, because the display does not explain the vertical references.

What matters is the **phase consensus**, not the amount. A rounded 0.0 mm or a missing rain field must not flip a symbol; the amounts are an additional plausibility check. Three conditions must hold simultaneously: the symbol contradicts the phase, the phase is supported by at least two models — the same threshold as the outlier filter of the vote —, and the displayed amounts are consistent with it. Towards rain this means at least 0.1 mm of rain and less than 0.1 cm of snow. Towards snow it means less than 0.1 mm of rain and, for a snow symbol, at least 0.5 cm of snow — the same amount that every other snow symbol requires (Section 14.3); otherwise the mapping would produce what the downgrade at the start of the cascade has rejected. Below that amount the liquid symbol does not remain either: "drizzle" next to zero millimetres of rain, at a site where the phase is snow, is precisely the contradiction the rule exists to resolve. The block then becomes cloudy — the same answer the snow downgrade gives without rain.

The mapping preserves the character of the precipitation:

Original code	**with solid phase**
Drizzle 51 / 53 / 55	Snowfall 71 / 73 / 75
Rain 61 / 63 / 65	Snowfall 71 / 73 / 75
Rain showers 80	Snow showers 85
Rain showers 81 / 82	Snow showers 86

Table 19: Phase mapping. Rain showers have three intensity levels, snow showers only two — the upper two therefore merge. The reverse direction applies correspondingly.

Freezing drizzle (56/57) and freezing rain (66/67) remain untouched. They carry a hazard statement of their own — a warm layer above a cold one, supercooled droplets — and are not an intensity variant of drizzle or rain. A phase determination from temperature, freezing level and amounts is not necessarily better than the model's diagnosis for such profiles. This holds where the diagnosis describes the site: for a site more than 300 m above its grid, a model diagnoses its own terrain, not the site. Its freezing code is read as the corresponding liquid one before the vote (66 as 61, 67 as 65, 56 as 51, 57 as 55), and the phase rule then decides as for any rain code — the same offset rule as for the ECMWF precipitation type (Section 6.1). The same applies in any hour in which the model's 2-m temperature at the site lies above 0 °C: without frost at the ground nothing freezes on contact.

The rule stands at the end of the cascade and not at its beginning: a shower downgraded because of too small an amount should not subsequently become a snow symbol.

Amount and coverage are two separate criteria. The amount thresholds of the downgrade rules test the weighted mean over the models that report any precipitation at all — that is, how much those who see some expect. How broadly the model group supports this is tested by the rules as a separate condition.

Coverage is a **weight fraction**, not a number of models: it relates the sum of weights of the models with precipitation to the sum of weights of those that supply any value for the block. Throughout

the method the weights stand for a model's skill in the Alpine region (Section 4.1); ignoring them at this point would be inconsistent — two high-resolution Alpine models carry a consensus differently from two coarse ones. Models without a value drop out of numerator **and** denominator: in the denominator alone they would effectively count as dry, and coverage would fluctuate with data availability instead of with the weather.

A mean over all models would be the wrong variable here: in it, the zeros of the dry models lower the value, and the rule would test the same uncertainty a second time. A single model would then have had to deliver five times the stated threshold for its symbol to survive the cascade — and that precisely in the situation in which a single high-resolution model is worth most, the convective summer shower. If a symbol survives the cascade with a narrow model basis, this is recorded as provenance.

The upgrade rules, by contrast, work with the mean over all models. A downgrade to "cloudy" because of weakly supported precipitation is therefore not reversed by the subsequent sunshine rules as long as this mean amounts to at least 0.1 mm. The cascade prefers the more conservative depiction there: weakly supported precipitation remains visible even if the models expect plenty of sun. For the question whether the overall situation is still wet, the precipitation load of all models is the appropriate variable — otherwise a single wet model could block the upgrade for all.

The last upgrade rule tests the block total and knows nothing of the sun; it therefore acts at night just as by day. There is no sunshine-based shower upgrade: below the amount threshold it would only restore the weak, narrowly supported shower that the downgrade has just rejected, so that the downgrade would apply only in the absence of sun.

The test is made against the **displayed** blended mean, that is, against exactly the value that stands next to the symbol — an override against a different variable would miss the contradiction. There is no additional coverage threshold: it would again suppress the single-model shower that the cascade has deliberately left in place since amount and coverage were separated, and the displayed mean contains coverage anyway, because the zeros of the dry models lower it. All that is required is that at least one model has precipitation.

Variables that do not enter the display are carried along: the number of hours with precipitation, the precipitation amount during the hours in the most severe rank, the strongest single hour within them, and the number of supporting models.

The **amount** is explicitly not a condition for escalation. A dry thunderstorm produces lightning with very little precipitation and is nonetheless an event at the summit; conversely, a brief heavy shower does not justify a symbol for six hours. It is recorded so that it can later be decided on the basis of data whether it should play a role.

12 Further Variables

12.1 Sunshine Duration

The interface supplies seconds per hour, computed on the server side — for ICON from direct radiation with a threshold of 120 W m^{-2}, for the other models from the difference between global and diffuse radiation. It is summed over the block (maximum 6.0 h) and averaged with weights over the models; this mean feeds the symbol (below).

The display shows the median per hour. Each hour carries a bar as high as its minutes of sunshine, a full bar being a full hour. Whether the sun shines is a question of how many models see it, and the mean answered it wrongly: at Mont Blanc it put 1.0 h of sun next to 21.4 cm of new snow, because two of six models saw sun (0, 0, 0, 0, 2.4 and 4.0 h). The length of the day needs no rule of its own: there is no bar at night, and a November morning fills its block only partly.

A model that precipitates in an hour contributes no sunshine for that hour. The median alone did not suffice, because the models themselves report sun while they precipitate: at Mont Blanc, AROME gave a full hour of sunshine at 0.7 mm of snow in the same hour. Sunshine duration is defined by direct irradiance of at least 120 W m^{-2} on a plane normal to the sun (WMO, 2003).

For most models it is not computed within the model but split off from the hourly mean of global radiation; such a separation estimates the direct share from brightness and does not know the cloud, and at high, snow-covered sites it is bright. Operational services post-process their elements so that they agree with one another; the rule here is the simplest form of that, applied per model and hour with the same 0.1 mm threshold as everywhere else. In the four recordings, no hour in which at least four models precipitate keeps any sun; dry hours are unaffected. Sun between showers remains where most models are dry and see it. The rule applies to the display only; the mean that feeds the symbol is unchanged.

The rule is a product heuristic and is not validated. A short shower and sunshine in the same hour are both real in showery weather, and an hourly sum says nothing about how long it rained within the hour; the rule removes such hours entirely. Equally, the median is not shown to be more accurate than the mean — it answers what is typical for the group of models and suppresses minority scenarios. The recordings document how both rules behave, not whether they improve the forecast. That the display uses the corrected median while the symbol keeps the uncorrected mean is a known asymmetry to be checked against visible counter-examples.

The value applies to the model grid point and is not converted to the site elevation. Unlike temperature or pressure, there is no admissible transformation for this: the sunshine duration at a summit depends on the horizon, on shading by neighbouring ridges, on slope angle and aspect, and on the vertical layering of the cloud. The direction of the error is not even consistent — a summit above a layer of high fog receives more sun than its grid point, a summit in the shadow of a neighbouring ridge less.

The variable enters the correction cascade in one place (Section 11.3): it upgrades a dry symbol to “mostly sunny”. It never restores a precipitation symbol that the cascade has rejected — neither a rain nor a snow shower — because a threshold that applied only under closed cloud would be no threshold. By construction, therefore, no precipitation symbol depends on the sunshine duration. A sensitivity analysis at the fixtures shows how much the dry symbols depend on it: with the sunshine duration set once to zero and once to the full block length, the symbol changes in 10 of 40 forecast blocks at the Zugspitze and in 4 of 40 at the Grossglockner — in every case between cloudy and mostly sunny.

The fair-weather symbol at high sites therefore rests on a grid-point quantity whose transfer to the summit is untested. Withdrawing the rule there would be a correction without evidence; the finding rests on two sites and one time window with related weather situations and is not a general error rate.

12.2 Wind Chill

The temperature felt in the wind is computed per hour from the displayed temperature T (°C) and the displayed mean wind V (km/h), with the wind chill index introduced jointly by the Canadian and the US weather services in 2001 (Osczevski and Bluestein[25]):

$$T_{\mathrm{wc}} = 13.12 + 0.6215T - 11.37V^{0.16} + 0.3965TV^{0.16}$$

The index describes the heat loss of exposed facial skin. It is defined up to 10 °C and from 4.8 km/h; outside this range there is no value. The index is formulated for the wind at 10 m; the displayed mean wind is used instead — at a summit the pressure-level wind with the exposure adjustment (Section 8.3), that is, the wind one stands in there.

The displayed mean wind at summits lies well above measurement (Table 26): at the Zugspitze 2.3 times the station value. The wind chill is therefore too cold there — at −10 °C and 30 km/h measured, the index gives −19.5 °C, with the displayed wind −23.3 °C. For this reason it is computed but not yet displayed: a number would suggest a precision the wind does not have. It will be shown once the mean wind at summits is corrected.

The provider's apparent temperature after Steadman [18] is not used for this: its derivation is not documented, and it is computed with the wind of the grid point, at summits often more than a thousand metres lower (Table 28).

12.3 Precipitation Type

Only ECMWF supplies this variable (Table 3). Freezing rain (3), wet snow (6) and ice pellets (8) are evaluated, searched over all hours of the block, unless the site lies more than 300 m above the ECMWF grid. The flags enter the phase decision (Table 9).

12.4 Other Blended Values

Variable	Block	Blend
Precipitation probability	mean	unweighted mean
Snow depth	last value	weighted mean
CAPE	maximum	maximum
Apparent temperature [18]	minimum	minimum
Relative humidity	mean	only in the humidity gate (Section 10.3)

Table 20: Aggregation of the other variables. Snow depth is a state value and is therefore not summed. No blended display value is produced for relative humidity.

The total snow depth is not displayed: ICON EU ends after about five days, ICON D2 after two (Table 2), so that from day 6 onwards the value comes from GFS alone. Weighted averaging would present a single-model value as a blended value.

The precipitation probability, by contrast, is displayed — as the height of the precipitation bars — although it too has a narrow basis: only ECMWF, GFS and ICON EU provide it (Table 3), so the mean runs over at most three of the six models. The difference from snow depth lies in the variable itself — it is already an ensemble measure of the supplying model, namely the fraction of its members above 0.1 mm in the respective hour, and carries its uncertainty within the value. If none of the three supplies data, no bar is drawn instead of suggesting a certainty for which there is no data basis.

The precipitation row shows it per hour. Each hour carries one bar: its height is the probability, its colour the phase at the site (Section 6.2). The bar averages over every model with an ensemble that has data, irrespective of the model selection: the selection chooses deterministic runs for amount and phase, and a user who keeps only ICON D2 and AROME — neither with an ensemble — would otherwise see no probability at all.

No positive probability is hidden. A bar stands wherever the ensembles give a probability above zero. Its colour comes from the selected models where they carry an amount in that hour. Where none does — the ensembles see precipitation, the deterministic runs stay dry — the colour follows the wet-bulb temperature at the site through the same phase curve that otherwise splits the amount (Section 6.2). Hiding these hours would have made a positive ensemble statement invisible because of the deterministic selection; in the summer recordings this affected blocks of up to 51 %. No bar is drawn only where the probability is zero or missing, or where no wet-bulb temperature exists either.

The colour of a bar therefore combines an ensemble probability with a phase derived from deterministic runs or from the site temperature. It is a combined display, not a joint probability of rain and snow.

The probability is not a value for the summit: it does not respond to the requested elevation (Section 3.5) and describes the grid point, at the Zugspitze about 1600 m below the summit. Only the colour of the bar belongs to the site.

The computation runs over all sources that carry a probability. Amount and probability come from different sets of models: the amount from all models with a valid precipitation value, the probability from ICON EU, ECMWF and GFS. AROME, AIFS and ICON D2 contribute amounts, never a probability.

This separation is structural and not a fault. An intersection of the two sets — the models that carry amount **and** probability at the same time — is therefore not suitable as a condition: where AIFS, AROME or ICON D2 alone carry the amount, it is not occasionally empty but necessarily so. It would then suppress not a contradictory pairing but the only statement available at all. In five recordings this affects 19 blocks, among them 2.58 cm of snow at the Grossglockner and 0.6 mm next to 52.5 % at the Zugspitze.

A high percentage next to deterministically zero millimetres is not a contradiction but precisely the case for which the variable exists: spread in the ensemble with a dry main solution.

The probability is thus an independent ensemble statement of its own, structurally narrow model basis — not the probability of the amount shown next to it. This distinction is carried by the text, not by a block.

Whether both numbers rest on the same set of models is carried along as a diagnostic with two cases: a structural difference of basis if the amount is carried solely by models without a probability field, and a disagreement between the models if the probability models additionally report hardly any prospect of precipitation. The threshold for this is 5 % and is product logic, not a meteorological limit.

The probability is averaged unweighted. The model weights in Table 7 answer the question which deterministic amount is more plausible at a summit; there the fine resolution of ICON EU justifies its higher weight. An event probability answers a different question: each contributor already supplies the result of its own ensemble, and their averaging should be guided by probabilistic skill, not by grid spacing. With the amount weights, ICON EU would carry 44 % of the sum of weights; measured at one site, 19 of 25 blocks dropped as a result, by up to 19 percentage points, solely because

the strongest weight lay on the driest model. A weighting by number of members (ICON EU 39, ECMWF 50, GEFS 30) would not be a solution either: it governs the sampling resolution within an ensemble and is not a measure of skill. Equal weighting is thus the robust starting point as long as no calibration against observations is available.

For ICON EU the variable arrives by a detour. Our own Open-Meteo instance does not issue it under the model key itself, but only under the collective key of the ICON group. Because the instance holds only the EU ensemble and neither the D2 nor the global one, this group has only one stage there: checked over seven days, 133 of 133 values agree with what the public service issues under ICON EU. The key is thus a transport route, not the technical name of the value.

Two conditions safeguard this, and both are part of the method, not of a server configuration. A value is adopted only for hours in which ICON EU itself supplies data — that is the horizon of the current model run, whereas a fixed number of hours would shift with the run's age. Should the global ensemble ever be added, the group would continue beyond this horizon; at the public service that amounts to 59 hours, and precisely these are discarded. If the retrieval fails or the horizon is missing, ICON EU is dropped entirely, and ECMWF and GFS remain.

Below half a millimetre or centimetre no number is shown. The display threshold is 0.5 — above the trace limit below which national services no longer give figures either: the US National Weather Service records rain below 0.25 mm and snow below 0.25 cm as "trace" without a number, and tipping-bucket gauges resolve only from 0.1 to 0.2 mm. Below that, a cell such as "0.1 cm" is not a statement but spurious precision.

The cost has been checked: over 65 days there is not a single one on which all four blocks fell below 0.5 although the daily total reaches 1 cm or 1 mm — the small values never stand alone for a day on which something happens. Only from 1.0 does this flip. The app does not display a daily total, which is why this calculation was necessary.

12.5 Computed but Not Displayed

The processing produces some variables that currently have no consumer in the display. They are named in the data model instead of running along silently.

Variable	Reason
Agreeing and participating models per symbol	Consensus measure; the cell offers no room for it (Section 14)
Spread of the wind across the models	Candidate for a range display
Total snow depth	see above
CAPE, convective fraction (showers per precipitation)	Diagnostic for thunderstorm risk
Low, mid and high cloud cover	collected for a cloud-layer display that does not exist
Boundary-layer height	Thermals; in practice GFS only
Composition of the gust: number of contributing models, of raised models and of models without a gust field, plus the highest derived value	Provenance diagnostic (Section 8.4)
Provenance of the cloud-top point: number of votes, highest valid cloud top, spread of the valid ones	records whether a point comes from one run or from both (Section 10.1)
Convective cloud base	collected, not yet evaluated

Variable	Reason
Lifting condensation level: mean, minimum and hours below the site	Diagnostic; the number itself is not a cloud base (Section 9.2)
Number of models behind the freezing level and rain–snow line	both rest on the models with a valid sign change in their own profile and thus on a changing set of contributors (Section 5.2)
Unexplained fraction of the precipitation amount per elevation band	keeps the budget closed where a profile does not support a row height (Section 13.1)
Lightning potential, updraft	ICON D2 only; carried along without a threshold — see Section 16.18
Maximum and minimum temperature over the models	the curve shows the mean; a band read as noise (Section 4.3)
Apparent temperature	derived by the provider without documentation, with the wind of its grid point (Section 3.3); its own wind chill is computed instead but not yet displayed either (Section 12.2)
Rain–snow line from the profiles and frost on every elevation band	the displayed snowfall level is estimated from the site (Section 5.4.2); both feed no display
Phase per elevation band	see Section 13.1
Mean sunshine duration	feeds the symbol only (Section 12.1)
Disagreement of profile and site temperature for the freezing level	diagnostic; the profile is kept (Section 5.3)
Precipitation probability with the display rule of the former cell	the bar shows the probability without it (Section 12.3)

Table 21: Variables computed but not displayed.

13 Elevation Bands

Summit elevation	Bands
$>$ 3500 m	Summit, then full thousands downwards — at most 2000 m below the uppermost thousand and never below 1000 m
2500–3500 m	Summit, then 2500 m (2000 m instead if less than 500 m apart), then steps of a thousand as long as they lie at 1000 m or above
$<$ 2500 m	Summit and one band at 1500 m; if the summit lies less than 500 m above it, 1000 m instead

Table 22: Derivation of the elevation bands. Minimum distance from the summit 500 m, span of the additional bands in high-alpine terrain at most 2000 m.

13.1 Phase per Elevation Band

The display shows no elevation grid. The precipitation row belongs to the site (Section 12.3): the probability does not depend on elevation, and one row per band would repeat the same bars in different colours. Where snow turns into rain is stated by the rain–snow line (Section 5.4), which rests on the phase per band described here.

The phase per band follows the same transition zone as at the site (Section 6.2); what differs is only where temperature, humidity and pressure at the row elevation come from. For each model, row and hour, the following order of precedence applies:

Native profile The row elevation lies between two pressure levels that the model provides and that both lie above its orography (Section 3.6.3). Temperature and humidity are interpolated to it. This is computed model atmosphere at this elevation and therefore the first choice.

Site-adjusted surface fields The row elevation lies above the orography but below the lowest usable level. This gap is not rare: at Mont Blanc it extends for ICON EU from 2868 m to about 3100 m, and the summit lies above it, but the 3000-m row in the middle of it. The fields that the service converts to **exactly this elevation** then carry the row.

No statement The row elevation lies below the model's orography. This model draws no bar in this row.

A separate query is made for each elevation band. The service's elevation parameter applies per request; taking the value of the summit row for the 2000-m row would mean claiming that one elevation applies to another.

The surface fields are deliberately **not** joined with the pressure levels into a common profile. A linear bridge through the boundary layer would look smooth and would be spurious precision: the 2-m temperature is itself a diagnosis from the surface and the lowest model level, not an independent anchor for an interpolation to the next pressure level. This holds regardless of how far away that level lies — the length of the span does not change the fact that its lower end does not carry the quantity it would have to carry.

That different models take different routes in the same row is admissible — which one applies depends on how high the model's orography lies. What would not be admissible is to present different validity limits as equal information quality; the provenance is therefore counted per model and row.

A row is shown only if the phase could be derived in at least two thirds of the precipitation hours of the block. Validity itself is still decided per hour — whether a level lies above the orography is a physical question and not one of display. Only the display is stabilised: a bar that appears and disappears between adjacent columns is harder to read than one that is consistently absent.

The amount is the same in every row. No source provides precipitation resolved by elevation: the service converts temperature and humidity to the requested elevation, but not the amount. An elevation gradient would be an assumption about orographic lifting, shading and exposure that is neither measured nor validated. The bar height therefore remains the same across all rows; only the colour depends on elevation.

The same separation applies to the new-snow depth in centimetres: it remains a site variable and is not recomputed per row. A number of centimetres per elevation band would simultaneously claim the same water basis at every elevation, elevation-accurate crystal formation and a reliable vertical distribution of precipitation.

A missing bar next to others indicates a model without a statement for that elevation; a colour of its own for this would be a phase that does not exist.

13.2 Which Bands Are Shown

Table 22 derives the bands from the site elevation. Of these, only what the models support is shown — and this is measured against the basis that **can exist**, not against all six models.

The difference is decisive. A model whose terrain lies above a band cannot represent it; that is a permanent property of its cell and not a disagreement about the weather. A fraction of the total weight would therefore take away votes from the row that never existed. At the Eiger (3967 m), ECMWF, AIFS and GFS fully carry the 3000-m band, whereas ICON EU, ICON D2 and AROME hold terrain above it there — 4.0 of 10.0 weight points, and at the same time the entire basis that can exist at this elevation. The row carries an amount in every block with precipitation, and the rain–snow line lies on average about 600 m below the summit.

An additional band therefore appears when two conditions coincide:

The possible basis is exhausted all models that can represent the elevation carry it — or at least three quarters of their weight.

At least two models carry it the same number as in the weather-code vote and the phase reconciliation, and for the same reason — a single model opinion should not determine anything on its own. At the Dufourspitze only GFS has its terrain below 3000 m; the row would otherwise be based there on a coarse global model alone, in the same picture as a row based on six.

Models, not approaches, are counted here: ICON EU and ICON D2 count as two, although D2 obtains its boundary conditions from the parent system. Whether a row ever depends on this pair alone is carried in the data as a separate number per band; at thirteen Alpine sites examined, the case does not occur.

A model without determinable orography remains in the reference set. It counts as not carrying, not as absent — otherwise a retrieval problem would reverse its meaning and increase coverage precisely through its failure. For the same reason, only the site row remains if the orography is entirely unknown: nothing is then established for any band, and ignorance is no evidence.

Site	derived	shown
Zugspitze, 2962 m	2962 / 2000 / 1000	2962 / 2000
Grossglockner, 3798 m	3798 / 3000 / 2000 / 1000	3798 / 3000
Eiger, 3967 m	3967 / 3000 / 2000 / 1000	3967 / 3000
Mont Blanc, 4808 m	4808 / 4000 / 3000 / 2000	4808 / 4000 / 3000
Dufourspitze, 4634 m	4634 / 4000 / 3000 / 2000	4634 / 4000
Chamonix, 1035 m	1035	1035

Table 23: Effect of the coverage rule. Removed are the bands that no model can represent and those for which only a single one would remain.

The two reference quantities answer different questions and are not mixed:

Against the suitable models Is the possible basis exhausted? **This quantity decides visibility.** Its denominator depends on model, location, elevation and the orography rule — not on response errors.

Against all models How broadly does the row stand compared with everything that exists? The denominator is fixed. The quantity is carried along as a diagnostic and decides nothing; at the Eiger at 3000 m it is 0.40 with the possible basis fully exhausted.

That a row rests on a narrower basis thus remains true and remains recorded. Omitting it altogether would be the stronger claim — it would say that the elevation cannot be represented, and that holds only if no model can represent it.

Four states summarise this: **complete** (all suitable models carry it), **sufficient** (at least three quarters of the possible weight), **incomplete** (below that) and **not representable** (no model can

represent the elevation). The first two show the row, the last two do not. The fourth is not a deficiency but a limit of the model chain; “incomplete” would mean that something is missing.

The three quarters are a prescribed display limit, not a physical quantity and not a quality threshold calibrated against observations. It prevents a row from appearing when a substantial part of the geometrically suitable basis fails. The usual case is the state **complete**: at thirteen Alpine sites examined, every visible additional row exhausts its possible basis to 100 %, so the limit only takes effect in the event of an actual failure. How often a row stands solely because of this limit, and which model is then missing, is carried in the data per band.

Which of the two visible states applies is not shown in the matrix — with cells of 40 × 28 pixels, any additional marking would cost more than it conveys. The state is carried per band in the forecast data and is described in this paper; it is not displayed.

The loss of a single model does not flip a row. If GFS drops out at Mont Blanc — with 1.0 the smallest weight —, 5.0 of the 6.0 possible weight points remain carrying; that is more than three quarters, and the rows below the summit remain. Such a failure is recorded as data incompleteness, not as a meteorological statement about the elevation.

Whether a band is carried at all is decided once per forecast run, not per time block. A model carries a band if it could derive a phase there in at least two thirds of the hours **of its own run** — not of the requested period, otherwise ICON D2, with its 48 hours against 312, would always fail. The weather situation thus determines the phase per band, but not which bands exist.

The computation runs over all supported models, not over those activated by the user. Otherwise the display selection would shift the topographic statement of the view. Which bars then stand in a row is, however, decided by the active models.

The application always decides the shape in this way, with the profile data of the run. Alongside it there is a fallback that derives the shape from the orographies alone and for this requires a fraction of the total weight — it does not know the suitable basis and therefore cannot use it as a reference. It thus shows at most fewer rows, never more, and is reserved for callers without run context.

The site row always remains. For it, the surface fields converted to its elevation are admissible even if the location lies below the model surface — in the Alpine region this is the normal case: in Chamonix, Zermatt, Innsbruck and Grindelwald the smoothed orography lies 99 to 1765 m above the actual location. The fields answer “what is likely at the requested location”; that is exactly what the service provides them for. What they do not answer is “how does the atmosphere change with elevation” — which is why they carry no further band below the orography.

There is thus never a permanently empty band in the data.

14 Marking of Estimated Values

Estimated variables — gusts without a usable model value (Table 14) — must be recognisable as such. Clive et al.[26] demonstrate for weather warnings that visual design variables measurably influence understanding and decision-making, and that extrinsic marking of uncertainty leads to more accurate judgements than intrinsic marking. The choice of marking is additionally constrained here by the cell geometry.

On mobile devices a cell measures 40 × 28 pixels. Many visual meanings are already inscribed in this grid: cell backgrounds for wind; the colour of the precipitation, sunshine and thunderstorm bars and of the temperature curve; text colours for snow, rain and freezing level; font weight for

warning and confidence — heavier for a gust from 80 km/h and for snow and rain amounts with a precipitation probability from 70 %; opacity for past values; column background for the change of day; a coloured edge for the current day.

Colour, opacity and font weight are thus taken. What remains available is a character within the value, an additional row and text outside the grid. The marking is implemented by a prefixed tilde.

A display of the number of models per cell cannot be represented with this geometry and would be up to fifteen times redundant per column, since the model basis depends on the lead time and not on the variable displayed.

14.1 Why the Snow Amount Carries No Marking

A marking of the snow amount would have to indicate that at least one contributing value is derived from precipitation and temperature rather than reported by the model. At high sites this applies almost always: the models' grid points lie several hundred to over a thousand metres lower there (Section 16.10) and report rain. At the fixture sites the fraction is 0 % (Zugspitze, 2962 m) and 67.5 % (Grossglockner, 3798 m) — so it varies not with the weather situation but with the grid elevation of the site. At Mont Blanc (4808 m) not a single value would get by without it.

A marking that stands next to almost every value distinguishes nothing; conversely, its absence would make a value appear to be a direct model statement. Since every snow amount is derived from the repartitioning described in Section 16.10, the marking would not distinguish anything at all.

The provenance is carried along in the data — it belongs in the method and in this paper, not in a cell of 40 × 28 pixels. The gust marking is not affected by this; there it does distinguish.

14.2 Water Phase and Precipitation Type

Two questions that need to be kept apart. The **water phase** states where an amount belongs in the budget: liquid, solid or both. The **type** states what follows from it on the mountain. Freezing rain is liquid and yet not rain.

The phase is determined hourly from the wet-bulb temperature at the site. The type usually coincides with it — except in three situations that only ECMWF reports, because only ECMWF provides a `precipitation_type` field.

Freezing rain overrides the partitioning — unless the site lies more than 300 m above the ECMWF grid (Section 6.1). It is diagnosed from the vertical temperature profile: snow melts in a warm layer and continues to fall, supercooled, through a near-surface frost layer. A wet-bulb temperature at the ground cannot reconstruct this stratification — it sees frost and infers snow. Glaze ice and powder demand opposite decisions, and the difference is not a nuance. Further from the grid the diagnosis says nothing about the site: at the Zugspitze ECMWF reported freezing rain for terrain 1619 m below the summit, and there the wet-bulb temperature decides.

The override acts **hourly**, not over the entire block: a block can carry one hour of freezing rain and five hours of snow, and that is exactly how it then appears in the numbers.

Ice pellets yield no new-snow depth. They fall as a solid and remain in the water budget, but do not settle like new snow. The snow-to-liquid ratio describes the volume of snow crystals; applied to ice grains, it yields a depth that does not exist.

Wet snow remains snow and keeps its depth, but is carried as a type of its own.

All three are single-model statements — no other model in the selection provides the field. That does not make them worthless: averaging away a diagnosis based on the vertical profile because the

other five models do not know the variable would discard the only available information, precisely for a rare and consequential event. The provenance is therefore carried along, and the statement is not suppressed.

14.3 Symbol and Amount Answer Different Questions

The weather symbol comes from the vote on the model weather codes and describes the model's **grid point**. The amounts are computed by us for the site. The two may diverge, and that is not a contradiction: a brief shower is an event — it affects visibility and wetness — even if nothing measurable remains in the end. The symbol answers the question about the event, the cell the question about the deposit. A drizzle or rain symbol may therefore stand next to a dash, as long as the phase at the site does not contradict it (Section 11.3).

A snow symbol is held to the amount. It requires a blended snow amount of at least 0.5 cm — the threshold from which the cell shows a number — at every step of the cascade that can produce one: in the downgrade at its start, irrespective of the sunshine duration, in the block override, and in the phase reconciliation at its end (Section 11.3). A snow symbol on a summit is read as new snow on the route; next to a dash it would promise a deposit that the numbers do not show. A snow shower under breaking skies that leaves less than 0.5 cm is consequently shown as cloudy or fair. In the five recordings (Section 18.1) no snow symbol stands next to a dash.

14.4 When a Dash Replaces a Number

Bar and amount stand in the same column, but they do not answer the same question. The bar height is a probability per hour and comes solely from the models with an ensemble; AROME, AIFS and ICON D2 contribute amounts, but never a probability.

The displayed amount is a sum over the six hours of the block, averaged over **all** models with valid precipitation — including those that report zero. It thus answers the question that belongs next to an ensemble probability: how much is to be expected on average, not how much falls in those models that expect this particular phase. A mean over only the carriers of a phase raises a single opinion to the level of a consensus.

The basis is the site-adjusted partitioning, not the provider's raw field. `rain` and `snowfall` apply to the model orography (Section 3.5) and do not close the budget: in five recordings there were blocks in which three models had precipitation, the site logic classified it as liquid, and the rain row nevertheless showed zero. An unconditional mean over the raw fields would be no solution — at Mont Blanc in winter it would turn zero visible rain blocks into nine.

Compared with a mean over only the models that carry a phase, the large amounts are the same where all models carry that phase — 8.3 cm at the Grossglockner, 14.9 cm at the Zugspitze, 29.7 cm at Mont Blanc. Only amounts carried by few models come out smaller, because a single opinion is not raised to the level of a consensus.

A bar can stand above a dash. The bar needs one selected model with an amount; the row shows the weighted mean over all of them. If a single model sees 0.5 mm and the others remain dry, a bar appears while the mean remains below the display threshold.

The threshold tests the **rounded** value, because the cell rounds: 0.45 appears as "0.5" and therefore counts as visible. Tested against the raw value, a dash would remain there under a bar without the models disagreeing.

The dash there does not claim "nothing is falling" but "no visible blended amount". Both statements are true, and the bars remain: precisely at transitions, with local showers and uncertain frontal

timing, they are the early signal that a multi-model product should provide. Coupling them to the threshold of the row would take this information away from the cell.

How broadly the precipitation is supported is carried along as a diagnostic — the weight fraction of the models with an appreciable amount. It separates the two situations that look the same in the cell: a dash with broad consensus versus a dash next to a minority signal. **A marking of this state in the cell is missing and remains open** — it must neither place a further character next to the number nor weaken the bars, which in small cells are only a few pixels high anyway.

If a model without an ensemble reports a trace while the ensembles say zero, no bar is drawn: its height would be zero. The amount row may still show a number — both statements are correct in themselves and rest on different models.

15 Prescribed Parameters

The following values are fixed in the processing. They are listed here in full, because without them none of the computations described can be reproduced; their origin and calibration status are assessed in Section 16.5.

Parameter	Value	Section
Phase at the site		
Transition zone, fully solid	$T_w \leq 0.0$ °C	6.2
Transition zone, fully liquid	$T_w \geq 1.5$ °C	6.2
Lower limit of the approximation, temperature	−20 °C	5.6.1
Lower limit of the approximation, humidity	5 %	5.6.1
Largest height of the site above the grid for precipitation type and freezing codes	300 m	6.1, 11
Lapse rate of the estimated snowfall level and freezing level	0.65 K / 100 m	5.3, 5.4.2
Freezing level		
Share of the covering weight for a height in metres	60 %	5.3
Share of the covering weight for frost on every row	75 %	5.3
Snow-to-liquid ratio		
Tolerance of the column selection	50 m	7.2
End of the SLR line	−16 °C	7.2
SLR below that	NWS table breakpoints	7.2
Smallest ratio	7.2	7.2
Onset of the wind reduction	20 km/h	7.3
Full effect of the wind reduction	40 km/h	7.3
Strongest wind reduction	0.7	7.3
Wind and gusts		
Blending 10 m → 850 hPa	1200–1600 m	8.1
Blending 850 → 700 hPa	2300–2800 m	8.1
Exposure adjustment summit / col	1.4 / 1.3	8.3
Gust factor below 1400 m	1.2	8.4
Gust factor at 2500 m	1.3	8.4

Parameter	Value	Section
Gust factor from 4000 m	1.45	8.4
Cap of the stacked product	1.7	8.4
Condensation level and elevation bands		
Factor of the dew-point depression	125 m per degree	9
Step of the elevation bands	1000 m	13
Minimum distance from the summit	500 m	13
Span in high-alpine terrain	2000 m	13
Distance of a level from the model surface	100 m	3.6.3
Share of the suitable model weight for a band to be shown	75 %	13.2
Models required for a band to be shown	2	13.2
Share of the total weight in the orography-only fallback	60 %	13.2
Valid hours required for a band to be shown	2/3	13.1

Table 24: Prescribed parameters of the processing chain.

Parameter	Value	Section
Thunderstorm risk		
CAPE thresholds	200 / 500 J/kg	10.1
Convective inhibition threshold	25 J/kg	10.1
Lifted index thresholds	0 / −4	10.1
Convective cloud top, per run	7000 m	10.1
Adjustment for model thunderstorm	2 points	10.1
Models required for the adjustment	2	10.1
Upper limit of the score	8	10.1
Humidity gate: probability	20 %	10.3
Humidity gate: relative humidity	70 %	10.3
Symbol and correction cascade		
Duration for symbol escalation	2 h	11.2
the same from rank 60	1 h	11.2
Rank from which one hour suffices	60	11.2
Precipitation threshold of a model	0.1 mm	6.1, 11.3
Display threshold rain / snow	0.5 mm / 0.5 cm	12.3
Downgrade for light precipitation	0.5 mm	11.3
Associated model fraction	40 %	11.3
Downgrade for weak showers	1.0 mm	11.3
Associated model fraction	50 %	11.3
Minimum snow for the snow symbol	0.5 cm	11.3
Minimum number of supporting models	2	11.3
Models for the phase reconciliation	2	11.3
Sunshine fraction for fair weather	50 %	11.3
Condensation level above the site	1500 m	11.3

Parameter	Value	Section
Models for severe reports	2	11.2
Display		
Minimum sunshine for a bar, per hour	0.05 h	12.1
Precipitation of a model that removes its sunshine, per hour	0.1 mm	12.1
Day and night for the labels of the temperature curve	06–18 / 18–06 h	4.3
Minimum span of the temperature scale	4 °C	4.3
Score from which a thunderstorm bar is drawn	2	10.5
Lead-time reweighting		
From lead time	3 days	4.2
Weight GFS / ECMWF / AIFS	0.5 / 2.0 / 2.0	4.2

Table 25: Prescribed parameters of symbol generation, thunderstorm risk, display and reweighting.

16 Limitations and Open Issues

16.1 Point Accuracy versus Model Resolution

Values are output for a named point, with elevation correction, exposure adjustment and elevation-dependent gust factor. The grids used have a horizontal spacing between about 2 km and 28 km (Table 2).

Each model holds its own terrain elevation at the relevant grid point, and in high mountains it deviates considerably from the summit elevation — in the coarsely resolved global models by more than a thousand metres, because a grid point averages over 25 km and the summit disappears within it. A model's 10-m wind applies to this elevation, not to the summit.

The corrections bridge the difference and in doing so produce a point-specific precision that the data basis does not support; the result is a site-adjusted estimate, not an observation. Bugnard et al.[22] demonstrate systematic deviations in temperature and wind for Alpine valleys that lie at the level of the model physics.

16.2 Undocumented Input Variables

CAPE, CIN, lifted index, freezing level, snowfall and apparent temperature are derived by the data provider without documented derivation (Section 3.3). Differences between models or changes in the derivation can therefore not be ruled out. For CAPE there is the additional factor that reanalyses systematically underestimate the value (Taszarek et al.[24]) — a finding for ERA5 and MERRA-2 whose transferability to the output used here is open (Section 10.2).

Our own safeguards do not cover this case. They compare the processing against recorded responses of the interface and trigger as soon as our **own** method changes. If, on the other hand, the provider changes the derivation of an input variable, the responses themselves change — and the test remains silent. This would become noticeable only through symptoms: values that no longer fit the weather situation.

16.3 The Snow-to-Liquid Ratio Remains a Heuristic

In the cold branch below −16 °C, the curve follows the breakpoints of the operational table down to −24 °C (Section 7.2). A monotonically continued curve would lie considerably higher there — at −20 °C 30:1 versus 18, at −24 °C 30:1 versus 8.8.

What remains is more fundamental. **First**, in the warmer branch the curve still lies above the table, at −10 °C with 20 versus 12. This is the difference between two published methods; which one is closer to observation in the Alpine region has not been tested here.

Second, both ends are adopted, not calibrated. The breakpoints come from an American station network, the line from an analysis over the Northern Hemisphere; there is no Alpine training set for either.

Third, the input is a single temperature. A profile method weights the layers by their share of the dendritic growth zone; the curve used cannot distinguish between a deep, thick growth layer and a thin one at the same height as long as the column maximum is the same.

Nor can the deviation be measured as an error against a better method. Rosenow et al.[8] compared four operationally used SLR methods across five winter storms and found **no** clearly best or worst one: even the fixed 10:1 assumption provided the best forecast at at least 19 % of stations in three of the five cases, while the profile-based method after Cobb did not achieve the largest share of best forecasts in any case. The spread between the methods is thus larger than the lead of any single one.

The more recent random-forest methods (Veals et al.[9], Pletcher et al.[10]) do not need this profile; they manage with near-surface predictors and would thus be applicable to the available fields. Their margin over the operational curves is considerable (Section 2.1).

Two limitations remain. Both are trained on North American stations; there is no Alpine training set. And Pletcher et al. themselves name where their method is weak: at ratios above 20 in cold, calm air — that is, in the band between −10 and −19 °C, in which the curve used here also reaches its highest values, with 26 at −16 °C. Both methods are thus most uncertain precisely where little water turns into a lot of snow. The data-driven approach is the most concrete way out of the heuristic, but not a ready-made replacement for the critical range.

16.4 Limits of Direct Model Gusts and Model Spread

There is no cap on the displayed gust relative to the **displayed** mean wind. The two quantities belong to different heights and partly to different grid points (Section 16.1); a ratio of them is therefore not a gust factor, and an upper limit derived from it would not be a physical bound but an arbitrary one.

A **model-internal** comparison must be distinguished from this: the 10-m gust and the 10-m wind of the same model run refer to the same field, the same grid point and the same diagnostic convention. Their ratio is a genuine gust factor, and a bound on it would be admissible. It is not currently set: this would require knowing which gust definition and which averaging time each source model uses, and the chosen value would have to hold up against measured gusts at summit stations. Both are outstanding. Instead, the provenance is carried along (Table 14), so that conspicuous cases can be evaluated retrospectively.

From below, the display is capped: the product of gust and exposure factor must not exceed 1.7. This value is prescribed, not derived — it limits how much the stacked heuristic may claim, and leaves model gusts untouched.

Three decisions in the aggregation already limit the spread without such a bound. The model gusts enter as a weighted mean, not as a maximum — a single model thus no longer determines the display on its own. The exposure adjustment does not act on them, so that the turbulent component already contained in the model is not amplified a second time. And a model gust that is too low is raised rather than discarded, so that the averaging does not run over an upwardly censored subset.

From high mountains there is one measurement: Naaim-Bouvet et al.[27] evaluate one winter at the Col du Lac Blanc, 2700 m in the French Alps. The gust factor determined there — the ratio of the one-second peak to the ten-minute mean — scatters widely up to about 5.5 and depends on the inflow direction. The probability that it exceeds 2 is 7 % with northerly winds, 14 % with southerly and easterly winds and over 60 % with westerly winds; westerly winds, however, account for only 2 % of the measurements, as the authors explicitly point out.

Because of the shorter averaging time, this quantity is not directly comparable with the model gusts used here. It therefore serves neither as evidence for an upper limit nor as evidence against one — it merely shows that ratios above 2 occur in Alpine terrain.

Letson et al.[28] arrive at a second finding for complex terrain: gusts there are not only larger than over flat terrain, they also decorrelate spatially. The mean probability that a gust occurs simultaneously at two neighbouring masts is 0.27. A gust value thus applies even less to the neighbouring ridge than a mean wind would.

A robust upper limit requires comparison with measured gusts at summit stations; model comparisons alone are not suitable for this.

One such comparison exists, and it shows the present state rather than a calibration. Against the DWD station on the Zugspitze (Deutscher Wetterdienst, open data, 10-minute values), the recorded January window yields 52 six-hour blocks; the site is classified as a summit and all six models are active.

Quantity	**Display / station**
Mean wind, ratio of sums	2.31
Mean wind without the exposure adjustment	1.65
Gust, ratio of sums	1.24
Blocks with the gust too low	15 of 52
Blocks with a station gust from 60 km/h	30 of 52
of these, gust more than 20 % too low	0

Table 26: Displayed wind at the Zugspitze against the DWD station, 1 to 13 January 2026. Station gust: maximum of the 10-minute peaks in the block; station mean: mean of the 10-minute means.

Two errors largely cancel in the gust. The displayed mean wind lies 2.3 times above the station — 1.65 times even without the exposure adjustment —, while the displayed distance between mean wind and gust is far too small: at a summit from 1558 m the floor is the cap of 1.7 times the wind before the exposure adjustment, and the gust lies only 21 % above the displayed mean wind, less than at an unexposed point at 4000 m (45 %). At the station the block maximum of the gust lies 2.26 times above the block mean. Mean wind and floor can therefore only be changed together: lowering the mean wind alone would pull the floor down with it and leave the gusts well below the measured ones.

The gusts that are too low all fall in weak-wind blocks; from a station gust of 60 km/h, none lies more than a fifth too low. The comparison rests on one station, thirteen days and archived series

without lead time, and it sets a block maximum of the station against a gust from hourly values. Whether the excess of the pressure-level wind is a property of the models or of the station site requires a second summit station.

16.5 Uncalibrated Parameters

Parameter	Value	Origin
Gust factors	1.2 / 1.3 / 1.45	prescribed
Exposure adjustments	1.4 / 1.3	empirical values, wind only
Cap of the stacked heuristic	1.7	prescribed, order of magnitude assessed
Condensation-level factor	125	literature [23]
CAPE thresholds	200 / 500	prescribed, not tested against lightning data
SLR breakpoints	8 … 26	literature [7], untested below −10 °C
Model weights	2.0 / 1.5 / 1.0	prescribed
Cascade thresholds	0.1 … 1.0 mm	prescribed
Lapse rate of the estimates from the site	0.65 K / 100 m	ICAO standard atmosphere
Sunshine: median, no sun in a wet model-hour	—	product heuristic, untested

Table 27: Prescribed parameters and their origin.

No method is calibrated against observations. The classification as summit or col comes from OpenStreetMap and does not represent exposure: a forested pre-Alpine summit receives the same factor as a free-standing ridge.

For wind in complex terrain there are measurement campaigns (Bugnard et al.[22]) and modelling approaches (Le Toumelin et al.[21]) against which a calibration could be aligned.

The conditions to which such a calibration is bound can be named independently of this. For the gust, the comparison stations must lie at summit level, otherwise the evaluation again compares two heights with each other instead of two quantities at the same place. It must cover at least one full winter season: gust ratio and snow-to-liquid ratio shift systematically in winter, convective gusts recede, gradient-driven ones increase. And the fit must be tested against periods other than those from which it was derived — otherwise the test measures itself.

How coarse the classification is is shown by an analysis of the Icelandic station network: Ágústsson and Ólafsson[29] find the mean gust factor to be independent of stability but clearly dependent on wind speed and local relief. It decreases with increasing wind and with station elevation, and it reaches 1.6 and more as soon as an upwind mountain rises at least 200 m above the station and lies closer than ten times this height. What matters is therefore the relief in the inflow direction, not the label of a point.

The gust factor used here, by contrast, increases with elevation. It refers, however, to a different quantity — the upper-level wind of a pressure level rather than the measured station wind —, so that the two findings are not in direct contradiction. This has not been tested.

For the exposure adjustment, a path other than calibrating the two factors is emerging: terrain-based downscaling methods derive the acceleration from the elevation model instead of from a classification of the location. Forthofer et al.[30] compare a mass-conserving and an additionally momentum-conserving method with the homogeneous wind field of a coarsely resolved forecast model and test

both against observations. Such a method would tie the adjustment to the actual terrain shape — ridge position, col shape and inflow direction — instead of to the question whether OpenStreetMap lists the point as a summit.

16.6 Different Limit Variables in ICON

Alongside the interpolated freezing level, ICON supplies a wet-bulb diagnosis of the snowfall level. The two quantities can lie several hundred metres apart. Averaged together they would yield a quantity without a consistent meaning — and since ICON EU ends at about +118 h (Table 2), the displayed concept would change at the transition from day 5 to day 6, which would make the display jump without any change in the weather.

The displayed freezing level is therefore computed uniformly from the air temperature; the snowfall level is reserved for phase determination (Section 5.5). The numeric row thus does not change concept at the end of the range. Whether the separation yields the better result in the Alpine region has not been tested against observations.

For the rain–snow line, by contrast, a change of source remains possible: when ICON EU drops out of range, a model can switch from `snowfall_height` to one of the other sources (Section 5.4.1). Threshold and derivation then change together. Where both sources overlap, the offset is 53 to 94 m in the median (Section 5.4.1); the source is carried per model in the data, so that such a jump is recognisable as a change of source and is not read as a change in the weather.

16.7 Ranking of Fog

The placement of fog in Table 16 above rain but below rain showers is a product decision. Rationale: loss of visibility makes orientation and descent more difficult; moreover, rain is represented several times in the display (amount row, bars, probability), fog only in the symbol. Heavy rain (code 65) ranks above fog, since at high intensities hazards arise that mere visual impairment does not present.

16.8 Different Model Sets for Score and Symbol

The thunderstorm score is computed over all models with data, quantitative variables only over those activated by the user. Rationale: CIN and lifted index are instability variables without a link to resolution and may be included conservatively for a warning value; numbers read off by the user, by contrast, must not come from deselected models. Score and symbol can therefore diverge.

16.9 Limit of Phase Determination

Jennings et al.[13] find only negligible gains from machine learning when rain–snow separation relies on near-surface variables alone. A switch to location-dependent thresholds (Jennings et al.[11]) would face the same limitation — and Fontrodona-Bach et al.[12] show for the Northern Hemisphere that shifting the thresholds does not even consistently improve skill.

16.10 Vertical Reference of the Phase Partitioning

The data service converts temperature, dew point, humidity and surface pressure to the requested elevation, but not the partitioning of precipitation into a liquid and a solid fraction. This partitioning describes the grid elevation of the model. The method thus places side by side two quantities that describe different places.

The offset is not a peculiarity of individual points. Measured over six Alpine summits and six models — 36 combinations — the grid point lies below the summit in 34 cases, by 1170 m on average and

2732 m in the extreme. Model terrain smooths summits away, and almost without exception in one direction. The two exceptions are ICON D2 and AROME at the Jungfraujoch, where the grid reaches 8 and 53 m above the summit — there the smoothing raises the col between two higher ridges instead of removing a peak.

Summit	ICON EU	ICON D2	AROME	ECMWF	GFS	AIFS
Mont Blanc, 4808 m	−1940	−728	−921	−2732	−2492	−2732
Grossglockner, 3798 m	−1071	−689	−566	−1529	−1594	−1529
Jungfraujoch, 3466 m	−393	+8	+53	−667	−1096	−667
Sonnblick, 3106 m	−793	−412	−399	−1115	−1104	−1115
Zugspitze, 2962 m	−1158	−654	−728	−1619	−1507	−1619
Säntis, 2502 m	−1137	−586	−460	−1304	−1434	−1304

Table 28: Elevation of the model grid point relative to the summit, in metres, for the cell the processing uses: the one the service chooses for the summit elevation (Section 3.6.3). Measured on 14 November 2026. ECMWF and AIFS carry the same values because at each of the six points they hit the same cell of the 0.25° grid — the columns are not duplicated by mistake.

Counted over seven days at the Zugspitze, ICON EU and ICON D2 report a liquid fraction in 29 and 28 % of the precipitation amount respectively, although their own snowfall level lies at or below the summit. At Mont Blanc (4808 m), four of five models report no snow at all over ten days, with 30 to 51 mm of precipitation and minimum temperatures between −11 and −15 °C — only ICON D2, whose grid point at 4080 m lies close to the summit, reports any.

Against this stands the repartitioning (Section 6.2): the amount remains that of the model, its partitioning results from the wet-bulb temperature at the site. It is the method's only phase decision, not a quantity alongside the partitioning of the grid point.

What is improved is thus the **phase** at the site. The spatial accuracy of the forecast precipitation amount remains unaffected — it is the amount of the grid point, with its orography. How large this remaining share is, is shown in the following section.

For the elevation bands below the summit, the same separation applies once more, but on a different basis: there the phase is carried not by the wet-bulb temperature at the site but by the pressure-level profile at the respective band elevation (Section 13.1). The amount there too remains that of the grid point. The display carries no band; the precipitation row is the site's.

16.11 Weather-Independent Elevation Correction for Three Models

The data provider's elevation conversion follows two different methods (Table 5). For ECMWF, GFS and AIFS it is a fixed rate of 0.650 K per 100 m, independent of inversion, stratification and advection.

This affects the variable on which the computation is based: every snow amount is derived from the wet-bulb temperature at the site (Section 6.2), and its three inputs come from exactly this conversion. For half of the models the phase decision thus rests on a weather-independent assumption, over distances given by the grid offset (Table 28): 1619 m at the Zugspitze for ECMWF and AIFS, 2732 m at Mont Blanc. In the second case, 17.8 K result from the fixed rate alone.

Observational studies show that the near-surface temperature gradient varies considerably with time of day, season and weather situation, and that the common standard rate does not transfer in general. Precisely the situations that matter here — frontal passage, inversion, a moist boundary layer under dry air aloft — are the ones in which it holds least.

What does not follow from this. The comparison shows a methodological asymmetry, not its sign: whether the fixed rate yields worse phases than one derived from the model's vertical structure has not been tested. A robust constant can do less harm over 2000 m than a coarsely resolved vertical profile. The processing therefore treats all six models identically and unchanged; a distance threshold beyond which models would drop out of phase determination would, at the current state of knowledge, be a hidden calibration.

The question can be decided, namely in the planned validation against station data: the extrapolation distance and membership of the fixed or the time-varying group belong there as stratification features of their own, together with their interaction. If the phase error grows with distance more strongly for the fixed rate than for the other models, there is a robust basis for a correction. Without this measurement there is none.

A technical route is also open and has not been tried: temperature and geopotential height of the pressure levels are available for each model and are already used for the freezing level and the column maximum. From them the site temperature could be interpolated vertically instead of adopting the 2-m temperature corrected at a fixed rate. The supplied 2-m temperature, however, contains near-surface model physics that a pressure level in the free atmosphere does not have — at the summit this can be an advantage, in other situations a loss. The approach should be tested offline, not rolled out.

16.12 Estimates from the Site

The displayed snowfall level (Section 5.4.2) always, and the displayed freezing level wherever the profile finds no line (Section 5.3), are carried from a value at the site with the fixed lapse rate of 0.65 K per 100 m. Neither knows the stratification of the day. In an inversion both lie on the wrong side: the freezing level is placed below the site where warm air lies above it, and a single snowfall level cannot express rain above and snow below. In winter this is not an edge case — the freezing level is estimated in 46 and 52 of 52 blocks of the two winter recordings.

For ECMWF, GFS and AIFS the estimate rests on a doubled assumption. Their site temperature has itself been carried from the grid point with the same rate (Section 16.11); carrying it on to the 0 °C line with the same rate again places the line where the standard atmosphere would place it from the grid point. The estimate then adds no information about the atmosphere above the site beyond the grid-point temperature. Only ICON EU, ICON D2 and AROME supply a site temperature that responds to the weather in their own conversion. The weighted mean combines both groups.

Where the source of the freezing level changes between profile and estimate, the number may jump; the source is carried with every value, but not shown. The estimate cannot contradict the site temperature it starts from — the value of the curve for an hour, the midpoint of the models' range for a block; the price is that it also cannot correct it.

What would decide the question: comparing the estimated freezing level with radiosonde or summit-station profiles in winter, separately for the two groups of the elevation correction and for inversion days.

16.13 Display Choices Without Validation

Several decisions of the display rest on legibility, not on a verification against observations.

The temperature curve shows the weighted mean; the extremes of the models are no longer drawn (Section 4.3). A frost night forecast by a single model is thus not visible, although the aggregation rules carry the extremes precisely for that case. The sunshine display uses the median and removes sunshine from hours in which a model precipitates (Section 12.1); whether that improves the forecast

has not been tested, and in showery weather it removes real sunshine together with the contradictory one. The symbol keeps the uncorrected mean, so display and symbol may disagree. The precipitation bar combines an ensemble probability for the grid point with a phase at the site (Section 12.3); its height is not a statement about the summit.

16.14 Changing Contributor Set of the Probability

The precipitation probability averages over all models that supply one (Section 12.3) — at most ICON EU, ECMWF and GFS. Which of these carry data in a block varies. The displayed number is therefore not a multi-model quantity with consistently the same composition: a value can rest on three ensembles or on a single one.

The number of contributors is carried along but blocks nothing. Blocking with a single contributor would remove most of the remaining probabilistic information precisely at the edge of the forecast range — that is where ICON D2 and AROME run out. For a later test against observations, one, two and three contributors must be evaluated separately.

The same restraint applies to the set of models behind the amount: that it differs from the one behind the probability is, given this field availability, the normal case and is noted as a situation, not treated as an error.

In addition, the variable carries the grid-point reference: it does not respond to the requested elevation (Section 3.5) and is not calibrated against station data.

16.15 The Forecast Precipitation Amount

When the snow chain was compared against observation, a discrepancy came to light that lies not in the method but in its input.

For 4 to 13 January 2026, all six models forecast between 24.1 and 39.4 mm of precipitation at the Zugspitze. The German Weather Service station at the same location measured 68.7 mm.

Source	Precipitation
AROME	39.4 mm
ICON D2	36.5 mm
ICON EU	36.4 mm
ECMWF IFS	27.6 mm
AIFS	24.9 mm
GFS	24.1 mm
Station DWD 05792	**68.7 mm**

Table 29: Zugspitze, 4–13 January 2026. The difference demonstrates a considerable discrepancy between model and station values. Its decomposition into model error and measurement-related undercatch remains open without station-specific metadata and wind correction — with snow and wind a precipitation gauge tends to catch too little rather than too much, so the station value may itself be too low.

This limits what can be said about the absolute new-snow amount. The chain computes the phase at the site and the ratio on physical grounds — but on a water basis whose magnitude is not confirmed here. A single period at one site is not enough to derive a correction from it.

16.16 The Layer Above the Site

The phase is determined from the wet-bulb temperature **at** the site. A warm layer above can melt a falling crystal even if it is freezing at the site; the phase at the ground thus does not depend on the site value alone.

Damping at the freezing level (Section 7.4) would not be a suitable proxy for this: it knows neither the depth nor the temperature of a melting layer. A suitable one would require temperature and humidity on several levels above the site together with integrated layer energy — variables that are not available at the necessary vertical resolution and would not be calibrated for the Alpine region. The point remains open.

16.17 Unit Handling in the Chain

The conversion from millimetres to centimetres takes place at one named point (Section 7.1). It would be more robust to carry all precipitation and water-equivalent amounts in millimetres throughout, to convert only at output and to let the field names carry the unit: a mix-up would then no longer be possible but would be visible as a type error.

The changeover is outstanding. It is not the correction of an existing error — the conversion is correct where it is — but a safeguard against future ones. It therefore ranks behind the points in this chapter that concern a running method.

16.18 Lightning Potential and Updraft Without a Threshold

Two variables target the thunderstorm itself instead of inferring it from environmental conditions as the score does: lightning potential and updraft. Both are supplied by ICON D2 alone, and only for its range of about two days.

They are collected and carried along, but enter no score and no display. The reason is the lack of a reference: a threshold could only be tested against lightning data, and without these any set limit would be an assertion. In the available recordings — two summer sites, 48 hours each — the lightning potential at the Zugspitze is zero throughout, at the Grossglockner between zero and 0.8; the updraft reaches 3.0 and 3.6 respectively. That suffices to establish unit, availability and missing-value convention, but not to justify a threshold.

What would be required: at least one complete thunderstorm season with events and non-events, spatially matching lightning detections with a defined time and radius window, separate evaluation by lead time, time of day, elevation and weather situation, and skill scores that weigh hits against false alarms. Until then both variables remain raw material.

17 Worked Examples

Three variables, computed in full. The values come from the production code, not from a re-implementation; anyone wishing to reproduce them will find the sections with the rules in brackets.

17.1 Gust Without a Model Value

A summit at 3000 m, mean wind 45 km/h at summit level, no active model supplies a gust field (Section 8.4).

Step	Value	Source
Mean wind	45 km/h	multi-model blend, elevation-blended

Step	Value	Source
Gust factor $g(3000)$	1.35	Section 8.4
Exposure adjustment summit	1.4	Section 8.3
Product, uncapped	1.89	1.35×1.4
Stacking cap applies	1.7	Section 8.4
Displayed gust	76.5 km/h	45×1.7

Table 30: Conservative lower bound when no model gust is available.

In this case the stacking cap applies: gust factor and exposure adjustment together give 1.89, the computation uses 1.7. This is not a special case but the usual one — at a site classified as a summit, the product already exceeds the cap from 1558 m, at a col from 2577 m. Above these elevations the displayed lower bound is thus no longer the product of the two heuristics but the cap itself.

The value carries the marking for derived gusts (Section 14), because no model supplies an anchor.

17.2 New-Snow Amount

4 mm of precipitation with a column maximum of −6 °C and 25 km/h surface wind, at a site at 3000 m whose wet-bulb temperature is below zero (Section 7).

Step	Value	Source
Total precipitation	4.0 mm	model value
Solid fraction at $T_w < 0$ °C	1.00	Section 6.2
Water as snow	4.0 mm	4.0×1.00
Snow-to-liquid ratio at −6 °C	16	Section 7.2
Wind correction at 25 km/h	0.925	Section 7.3
Product	59.2 mm	$4.0 \times 16 \times 0.925$
Displayed amount	5.9 cm	$\frac{59.2}{10}$

Table 31: The division by ten is the conversion from millimetres to centimetres; the ratio itself is dimensionless (Section 7.1). The freezing level does not enter this computation — the reason is given in Section 7.4.

17.3 Phase at the Site

1.5 °C air temperature, 65 % relative humidity, 710 hPa surface pressure — the values of a summit at about 2900 m (Sections 5.6 and 6.2).

Step	Value	Source
Air temperature	1.5 °C	computed for site elevation
Wet-bulb temperature, psychrometric	−1.09 °C	Section 5.6
Solid fraction	1.00	Section 6.2

Table 32: The thermometer is above zero, the wet-bulb temperature below — the precipitation arrives entirely as snow.

For comparison: the same equation with sea-level pressure instead of the site's 710 hPa would give −0.61 °C, that is, 0.48 K higher. This is the pure pressure effect — at the phase transition about a hundred metres of elevation, depending on the temperature gradient.

Stull's direct formula, which applies without surface pressure, would arrive at −1.29 °C here. It thus lies not between the two pressure cases but below both: at this point, its own approximation error and the missing pressure reference act in the same direction.

This is the usual case, but not a rule. Sampled on a regular grid over Stull's stated domain — −20 to 50 °C in steps of 0.5 K, 5 to 99 % relative humidity in steps of 1 %, 13,395 points at 1013.25 hPa —, the approximation lies below the psychrometric computation in 59.5 % of points and above it in 40.5 %, by up to −1.47 and +3.98 K respectively. The distance between the approximation and the production value is thus composed of two components whose signs change independently of each other. It is therefore not suitable as a measure of the correctness of the production value (Section 5.6.2).

18 Reproducibility

The figures in this document come from two kinds of data. They differ in how far they can be reproduced, and the difference is stated here instead of at each figure.

18.1 Recorded Interface Responses

The site-specific figures for the Zugspitze, the Grossglockner and Mont Blanc come from interface responses recorded on 25 August 2026 from the application's own Open-Meteo instance and archived unchanged. Each covers 312 hourly steps: for the summer recordings three past days and ten forecast days around the recording date, for the winter recordings the stated past period.

Site	Requested point	Elevation	Period (local time)
Zugspitze	47.4211, 10.9853	2962 m	22 Aug – 3 Sep 2026
Zugspitze	47.4211, 10.9853	2962 m	1 – 13 Jan 2026
Zugspitze	47.4211, 10.9853	2962 m	5 – 17 Jan 2026
Grossglockner	47.0745, 12.6939	3798 m	22 Aug – 3 Sep 2026
Mont Blanc	45.8326, 6.8652	4808 m	1 – 13 Jan 2026

Table 33: Recorded responses. The unit of a figure is stated with it: a model hour is one valid hour of one model, a block one six-hour block of the multi-model blend.

The processing chain is tested automatically against these responses. One test compares the complete output with a stored reference and fails on any change of the method; others recompute figures of this paper directly, among them the offset of the rain–snow line at a change of source (Section 5.4.1), the dependence of the symbol on the sunshine duration (Section 12.1) and the comparison of Stull's approximation with the psychrometric solution (Section 17.3).

18.2 Live Queries

The remaining figures — among them those for twelve, thirteen and fourteen Alpine sites, the grid offsets in Table 28, the derived pressure levels and the valley sites — were measured on live interface responses at the time each part of the method was built, the derived-level check on 25 August 2026. They are not part of the recorded set. The interface does not retain past runs, so these figures can be repeated with the same procedure but not with the same data; a repetition will reproduce their order of magnitude, not their exact values.

References

[1] Y. Seity *et al.*, 'The AROME-France convective-scale operational model', *Monthly Weather Review*, vol. 139, no. 3, pp. 976–991, 2011, doi: 10.1175/2010MWR3425.1.

[2] G. Zängl, D. Reinert, P. Rípodas, and M. Baldauf, 'The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core', *Quarterly Journal of the Royal Meteorological Society*, vol. 141, no. 687, pp. 563–579, 2015, doi: 10.1002/qj.2378.

[3] S. Lang *et al.*, 'AIFS — ECMWF's data-driven forecasting system', *arXiv preprint*, 2024, doi: 10.48550/arXiv.2406.01465.

[4] L. Magnusson *et al.*, 'Skill of medium-range forecast models using the same initial conditions', *Bulletin of the American Meteorological Society*, vol. 103, no. 9, pp. E2050–E2068, 2022, doi: 10.1175/BAMS-D-21-0234.1.

[5] P. J. Roebber, S. L. Bruening, D. M. Schultz, and J. V. Cortinas Jr., 'Improving snowfall forecasting by diagnosing snow density', *Weather and Forecasting*, vol. 18, no. 2, pp. 264–287, 2003, doi: 10.1175/1520-0434(2003)018<0264:ISFBDS>2.0.CO;2.

[6] D. K. Cobb Jr. and J. S. Waldstreicher, 'A simple physically based snowfall algorithm', in *21st Conference on Weather Analysis and Forecasting / 17th Conference on Numerical Weather Prediction, Washington, DC*, American Meteorological Society, 2005, p. 2.

[7] M. S. Cohen, 'Snow/Liquid Ratio', 2013. [Online]. Available: https://www.weather.gov/media/mdl/SLR.pdf

[8] A. A. Rosenow, H. D. Reeves, and D. D. Tripp, 'Evaluation of ensemble snowfall forecasts using operationally used snow-to-liquid ratios in five winter storms', *Weather and Forecasting*, vol. 38, no. 10, pp. 2135–2147, 2023, doi: 10.1175/WAF-D-22-0202.1.

[9] P. G. Veals *et al.*, 'Predicting snow-to-liquid ratio in the mountains of the western United States', *Weather and Forecasting*, vol. 40, pp. 1825–1838, 2025, doi: 10.1175/WAF-D-24-0233.1.

[10] M. D. Pletcher *et al.*, 'Snow-to-liquid ratio prediction over the contiguous United States using machine learning', *Weather and Forecasting*, vol. 41, no. 2, pp. 317–338, 2026, doi: 10.1175/WAF-D-25-0124.1.

[11] K. S. Jennings, T. S. Winchell, B. Livneh, and N. P. Molotch, 'Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere', *Nature Communications*, vol. 9, p. 1148, 2018, doi: 10.1038/s41467-018-03629-7.

[12] A. Fontrodona-Bach, B. Schaefli, R. Woods, and J. R. Larsen, 'Estimating robust melt factors and temperature thresholds for snow modelling across the Northern Hemisphere', *Hydrology and Earth System Sciences*, vol. 30, pp. 2613–2636, 2026, doi: 10.5194/hess-30-2613-2026.

[13] K. S. Jennings *et al.*, 'Machine learning shows a limit to rain-snow partitioning accuracy when using near-surface meteorology', *Nature Communications*, vol. 16, no. 1, p. 2929, 2025, doi: 10.1038/s41467-025-58234-2.

[14] O. Brasseur, 'Development and application of a physical approach to estimating wind gusts', *Monthly Weather Review*, vol. 129, no. 1, pp. 5–25, 2001, doi: 10.1175/1520-0493(2001)129<0005:DAAOAP>2.0.CO;2.

[15] P. Zippenfenig, 'Open-Meteo.com weather API'. 2024. doi: 10.5281/zenodo.14582479.

[16] European Centre for Medium-Range Weather Forecasts, 'ECMWF's AI forecasts become operational'. 2025.

[17] 'Manual on Codes, Volume I.1, Part A — Alphanumeric Codes. Code table 4677: Present weather reported from a manned station', no. WMO-No. 306. Geneva, 2019.

[18] R. G. Steadman, 'A universal scale of apparent temperature', *Journal of Climate and Applied Meteorology*, vol. 23, no. 12, pp. 1674–1687, 1984, doi: 10.1175/1520-0450(1984)023<1674:AUSOAT>2.0.CO;2.

[19] R. Stull, 'Wet-Bulb Temperature from Relative Humidity and Air Temperature', *Journal of Applied Meteorology and Climatology*, vol. 50, no. 11, pp. 2267–2269, 2011, doi: 10.1175/JAMC-D-11-0143.1.

[20] K. G. Libbrecht, 'The physics of snow crystals', *Reports on Progress in Physics*, vol. 68, no. 4, pp. 855–895, 2005, doi: 10.1088/0034-4885/68/4/R03.

[21] L. Le Toumelin, I. Gouttevin, N. Helbig, C. Galiez, M. Roux, and F. Karbou, 'Emulating the adaptation of wind fields to complex terrain with deep learning', *Artificial Intelligence for the Earth Systems*, vol. 2, no. 1, p. e220034, 2023, doi: 10.1175/AIES-D-22-0034.1.

[22] A. Bugnard, M. Collaud Coen, M. Hervo, D. Leuenberger, M. Arpagaus, and S. Monhart, 'Comparison of temperature and wind profiles between ground-based remote sensing observations and numerical weather prediction model in complex Alpine topography: the Meiringen campaign', *Atmospheric Measurement Techniques*, vol. 18, pp. 1039–1061, 2025, doi: 10.5194/amt-18-1039-2025.

[23] M. G. Lawrence, 'The relationship between relative humidity and the dewpoint temperature in moist air: a simple conversion and applications', *Bulletin of the American Meteorological Society*, vol. 86, no. 2, pp. 225–234, 2005, doi: 10.1175/BAMS-86-2-225.

[24] M. Taszarek, N. Pilguj, J. T. Allen, V. Gensini, H. E. Brooks, and P. Szuster, 'Comparison of convective parameters derived from ERA5 and MERRA-2 with rawinsonde data over Europe and North America', *Journal of Climate*, vol. 34, no. 8, pp. 3211–3237, 2021, doi: 10.1175/JCLI-D-20-0484.1.

[25] R. Osczevski and M. Bluestein, 'The new wind chill equivalent temperature chart', *Bulletin of the American Meteorological Society*, vol. 86, no. 10, pp. 1453–1458, 2005, doi: 10.1175/BAMS-86-10-1453.

[26] M. A. T. Clive, E. E. H. Doyle, S. H. Potter, C. Noble, and D. M. Johnston, 'How visual design of severe weather outlooks can affect communication and decision-making', *Weather, Climate, and Society*, vol. 15, no. 4, pp. 979–997, 2023, doi: 10.1175/WCAS-D-23-0010.1.

[27] F. Naaim-Bouvet, M. Naaim, H. Bellot, and K. Nishimura, 'Wind and drifting-snow gust factor in an Alpine context', *Annals of Glaciology*, vol. 52, no. 58, pp. 223–230, 2011, doi: 10.3189/172756411797252112.

[28] F. Letson, R. J. Barthelmie, W. Hu, and S. C. Pryor, 'Characterizing wind gusts in complex terrain', *Atmospheric Chemistry and Physics*, vol. 19, no. 6, pp. 3797–3819, 2019, doi: 10.5194/acp-19-3797-2019.

[29] H. Ágústsson and H. Ólafsson, 'Mean gust factors in complex terrain', *Meteorologische Zeitschrift*, vol. 13, no. 2, pp. 149–155, 2004, doi: 10.1127/0941-2948/2004/0013-0149.

[30] J. M. Forthofer, B. W. Butler, and N. S. Wagenbrenner, 'A comparison of three approaches for simulating fine-scale surface winds in support of wildland fire management. Part I. Model formulation and comparison against measurements', *International Journal of Wildland Fire*, vol. 23, no. 7, pp. 969–981, 2014, doi: 10.1071/WF12089.