Accelerating Earth System Model Physics Development with Agentic Coding

Adrian Hill[1] & Paul Field[2]

(1)Research Department, ECMWF, Reading, United Kingdom and (2) Met Office, Exeter, United Kingdom

1. Motivation and objectives

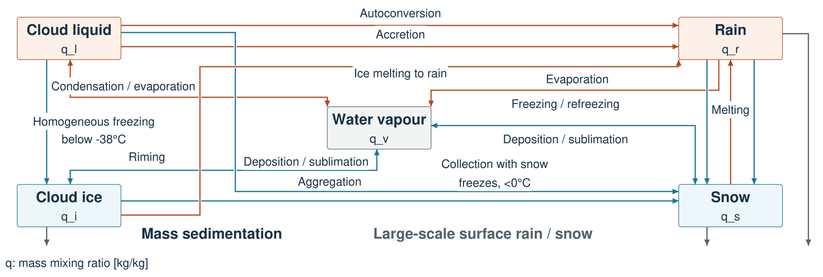
- Assessing the scientific and operational benefits and computational cost of new prognostic science schemes in Earth system models requires implementation, testing and validation. This process can take years.
- Agentic coding could substantially shorten the development and testing cycle for new schemes, enabling earlier assessment of whether they improve or degrade model performance and merit consideration for operational use.
- Here we use the implementation of the multi-moment Met Office Cloud AeroSol Interacting Microphysics (CASIM) scheme in OpenIFS as a test case to understand the strengths and limitations of agentic coding.
- **The aim is to evaluate how efficiently AI agents implement CASIM in OpenIFS and identify the technical and scientific issues encountered during implementation.**

2. OpenIFS and cloud microphysics

OpenIFS is ECMWF's open-source global forecast-only* IFS model for NWP, climate research, and education (see https://github.com/ecmwf-ifs/openifs for code and details).

OpenIFS/IFS uses single-moment **CLOUDSC** for large-scale mixed-phase clouds and precipitation, predicting cloud fraction and mass mixing ratios (q) for cloud liquid, cloud ice, rain and snow with prescribed particle number concentrations.

IFS/OpenIFS Cy49r1: CLOUDSC: large-scale microphysics



q: mass mixing ratio [kg/kg]

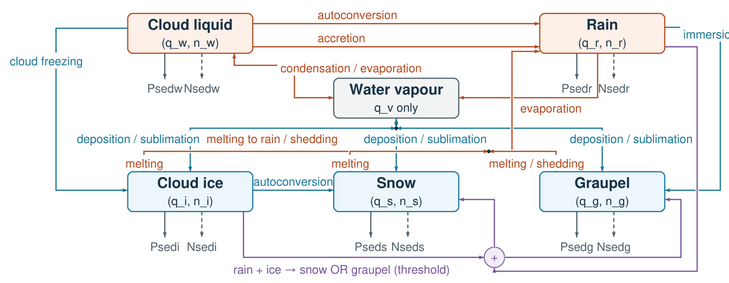
***What does OpenIFS include/exclude?**

- Includes the atmospheric model, land, waves, hydrostatic core and composition.
- Excludes DA, the IFS LAM configuration, observation processing and the non-hydrostatic core.

3. CASIM

CASIM (Cloud AeroSol Interacting Microphysics), the Met Office's open-source, configurable multi-moment bulk scheme, represents clouds, precipitation and aerosol–cloud interactions across model scales (https://github.com/MetOffice/casim).

Operational in the Met Office regional model since 2026, double-moment (2M) CASIM predicts mass mixing ratio and number concentration for cloud liquid, rain, cloud ice, snow and graupel.



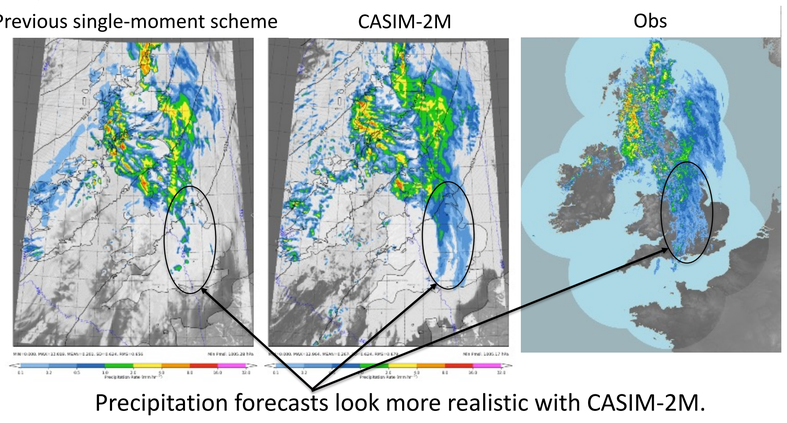
q: mass mixing ratio [kg/kg] n: number mixing ratio [particles/kg]

Paired downward arrows: Psedx uses mass-weighted; Nsedx number-weighted fall speeds → size sorting.

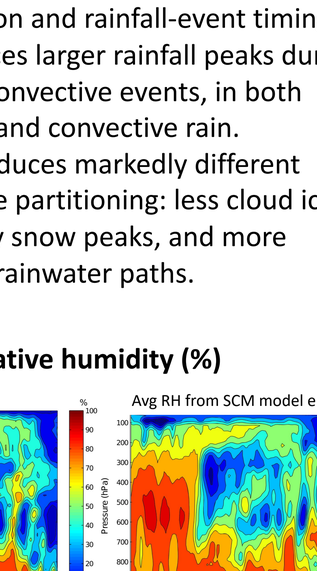
CASIM options include a third moment for rain, snow and graupel, and cloud–aerosol interactions (Field et al., 2023).

Operational Met Office Unified Model forecasts

Previous single-moment scheme | CASIM-2M | Obs



Precipitation forecasts look more realistic with CASIM-2M.

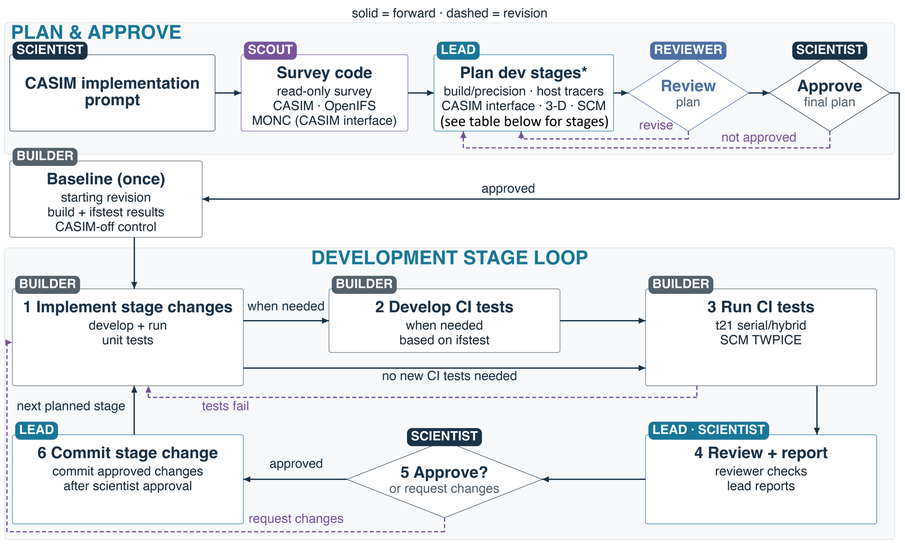


CASIM-2M better predicts how often light and very heavy rain occur in the regional model.

CASIM was chosen as a test case because of the authors' familiarity with the scheme. Other physics schemes could provide further tests of the agentic coding approach.

4. Agentic CASIM integration in OpenIFS

Aim: AI agents plan, code and test CASIM integration in OpenIFS.
Humans review and approve. AI writes all code.



- The workflow above was initially applied to OpenIFS 48r1
  - ***Implementation time ~ 1 week***
- Workflow then used to upgrade to OpenIFS 49r1
  - ***Upgrade time ~ 1 day***

Phase	Stage	One-line definition
0	**Build system**	Integrate CASIM and fallback stubs into the OpenIFS build, establishing a CASIM-off baseline.
1	**Field registration**	Register hydrometeor mass, number and higher-moment fields, their configuration switches and output identifiers.
2	**Convective transport**	Extend convection to transport CASIM particle-number and higher-moment fields alongside hydrometeor mass.
3	**Initialisation**	Create yocasim.F90 to manage global and per-thread CASIM initialisation.
4	**Coupling interface**	Translate OpenIFS state and forcing into CASIM inputs, then return microphysics tendencies and precipitation fluxes.
5	**Physics dispatch**	Modify CALLPAR to select CASIM instead of CLOUDSC and incorporate its results while preserving host tendency diagnostics.
6	**Downstream coupling, diagnostics and output**	Check radiation, convection and advection compatibility, and provide hydrometeor mass, number and higher-moment diagnostics through GRIB output.
7	**Cloud-fraction and precipitation-cover bridge**	Supply the host cloud-fraction update and address precipitation-cover diagnostics needed by downstream physics.
8	**Single-column verification**	Test transport, warm-rain and full microphysics using field output, water-path and precipitation diagnostics, plus backward-compatibility checks.
9 (linked)	**Convective precipitation-number sources**	Trace convective rain/snow mass pathways with diagnostics and, where required, add and validate consistent particle-number sources.

* Stages defined in "Plan dev stages" step of Plan & Approve workflow

Summary of implementation prompt

1. Integrate CASIM bulk microphysics into OpenIFS as a switchable alternative to CLOUDSC, supporting both the full atmospheric model and the single-column model.
2. Reuse mass/vapour fields. Add number and optional higher moments with consistent initialisation, transport and convection coupling.
3. Initially omit aerosol activation. Use host droplet number as an activation target, allowing prognostic number to evolve.
4. Provide a thread-safe interface with correct units, vertical ordering and complete flux profiles for downstream physics.
5. Preserve standard diagnostics, add CASIM diagnostics and cross-check precipitation. Add convective number sources only where justified.
6. Implement in stages. Require unchanged CASIM-off results and separate CASIM-active tests. Software tests alone cannot establish scientific equivalence.

5. OpenIFS SCM: CASIM vs CLOUDSC

Comparison of CASIM and the standard CLOUDSC microphysics in the OpenIFS single-column model (SCM) using a deep convective model intercomparison case, TWP-ICE (Davies et al., 2013).

Initial analysis

- CASIM broadly reproduces CLOUDSC's RH evolution and rainfall-event timing but produces larger rainfall peaks during the main convective events, in both stratiform and convective rain.
- CASIM produces markedly different condensate partitioning: less cloud ice, larger early snow peaks, and more persistent rainwater paths.

Relative humidity (%)

6. Summary

The CASIM implementation in OpenIFS demonstrates the effectiveness of agentic coding when expert guidance and reference code are available.

- **Rapid implementation enabled scientific testing:** Agents wrote all integration code, achieving successful SCM and coarse global simulations in approximately one week in OpenIFS 48r1, followed by an upgrade to 49r1 in one day. For context, implementing CASIM in the Met Office Unified Model took approximately five FTE-years over ten years although this effort also included CASIM development, so the efforts are not directly comparable. Agentic coding rapidly made OpenIFS–CASIM available for initial SCM comparisons with CLOUDSC, enabling earlier decisions on whether to proceed with broader, more detailed scientific testing.
- **Prior experience supported this speed:** Previous end-to-end implementation without agents provided knowledge of the technical and scientific pitfalls, helping guide the agents and assess their output. Existing implementations supplied reference code and coupling examples.
- **Expert review remained essential:** For example, agents proposed an incorrect number-concentration correction and struggled with condensation and cloud-fraction coupling, demonstrating the need for expert review. Whether comparable speed and scientific reliability are achievable for a new implementation from scratch, particularly with less domain expertise, remains an open question.

References

Davies, L., et al. (2013), A single-column model ensemble approach applied to the TWP-ICE experiment, *J. Geophys. Res. Atmos.*, 118, 6544–6563, doi:10.1002/jgrd.50450.

Field, P.R., Hill, A., Shipway, B., Furtado, K., Wilkinson, J., Miltenberger, A., et al. (2023) Implementation of a double moment cloud microphysics scheme in the UK Met Office regional numerical weather prediction model. *Quarterly Journal of the Royal Meteorological Society*, 149(752), 703–739. Available from: https://doi.org/10.1002/qj.4414