Plans for surface processes and surface data assimilation in HARMONIE-AROME

Patrick Samuelsson
SMHI

with acknowledgements to colleagues within HIRLAM, at Météo-France, MetCoOp

The ALADIN-HIRLAM cooperation

… uses the ALADIN-HIRLAM NWP system which includes three different configurations:

- **HARMONIE-AROME**
- **AROME-France**
- **ALARO**

All of them relate to the model SURFEX for surface processes.

http://www.umr-cnrm.fr/aladin/IMG/pdf/nl10.pdf

The ALADIN-HIRLAM cooperation

… uses the ALADIN-HIRLAM NWP system which includes three different configurations:
- **HARMONIE-AROME**
- **AROME-France**
- **ALARO**

All of them relate to the model SURFEX for surface processes.

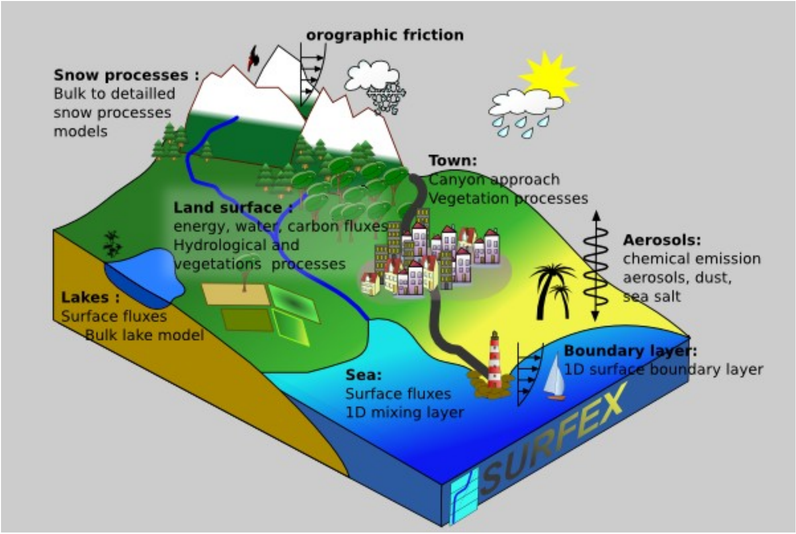
Operational cooperation among HIRLAM countries:
- **MetCoOp (Sweden, Norway, Finland) runs HARMONIE-AROME in ensemble system (10 members). Negotiations are ongoing with Estonia and Lithuania to join.**
- **Iceland, Denmark, Ireland and the Netherlands have the intention to form an operational collaboration.**

http://www.umr-cnrm.fr/aladin/IMG/pdf/nl10.pdf

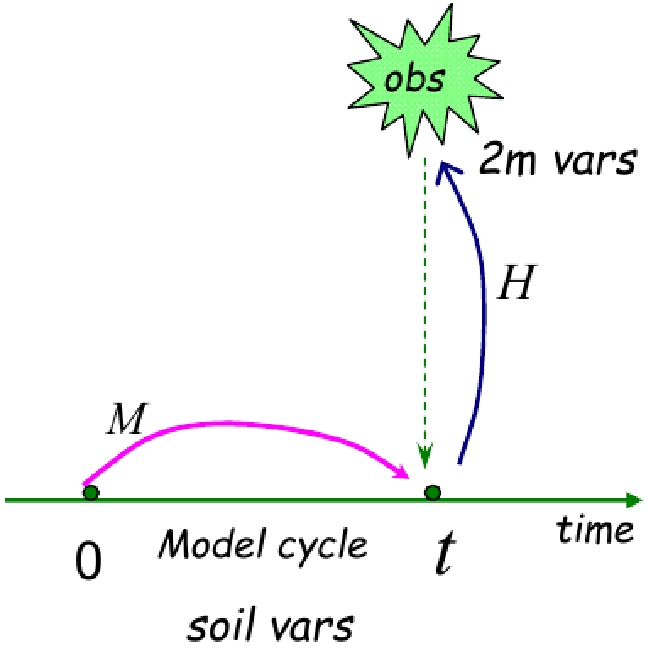
The three main components of surface modelling for NWP

Physiography

Processes



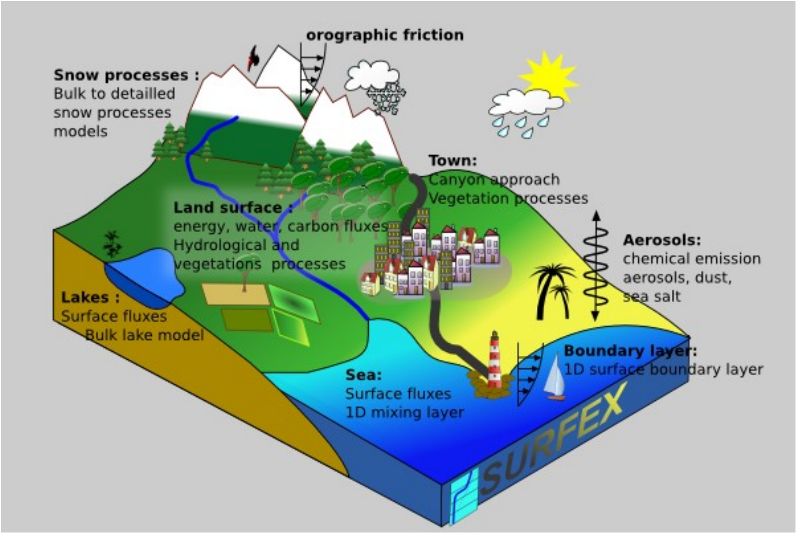
Assimilation



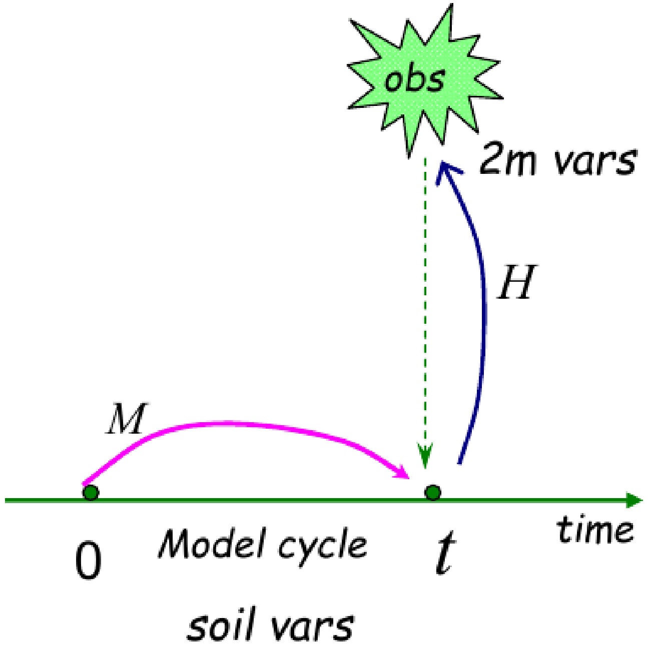
The three main components of surface modelling for NWP

Physiography

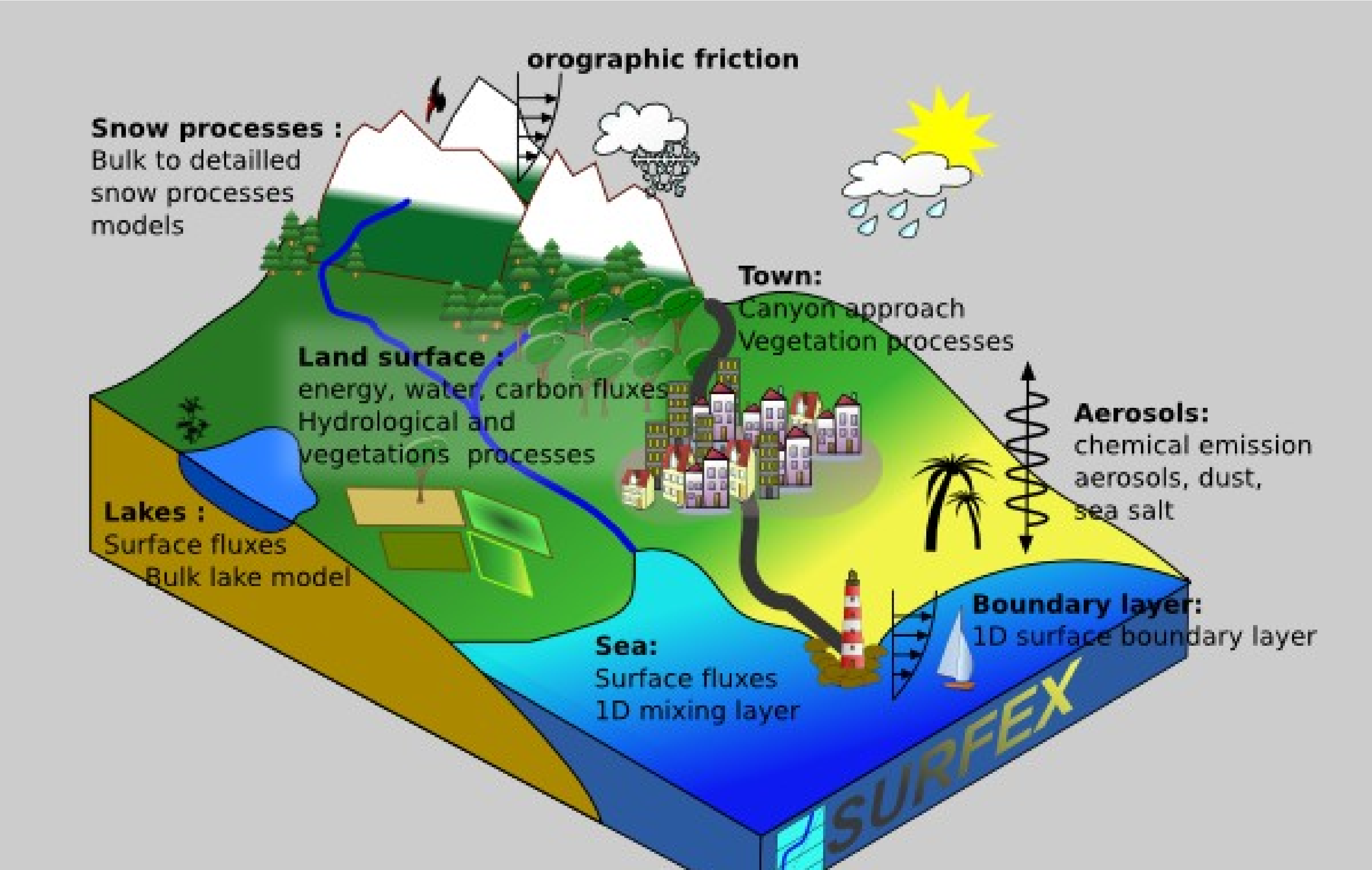
Processes



Assimilation

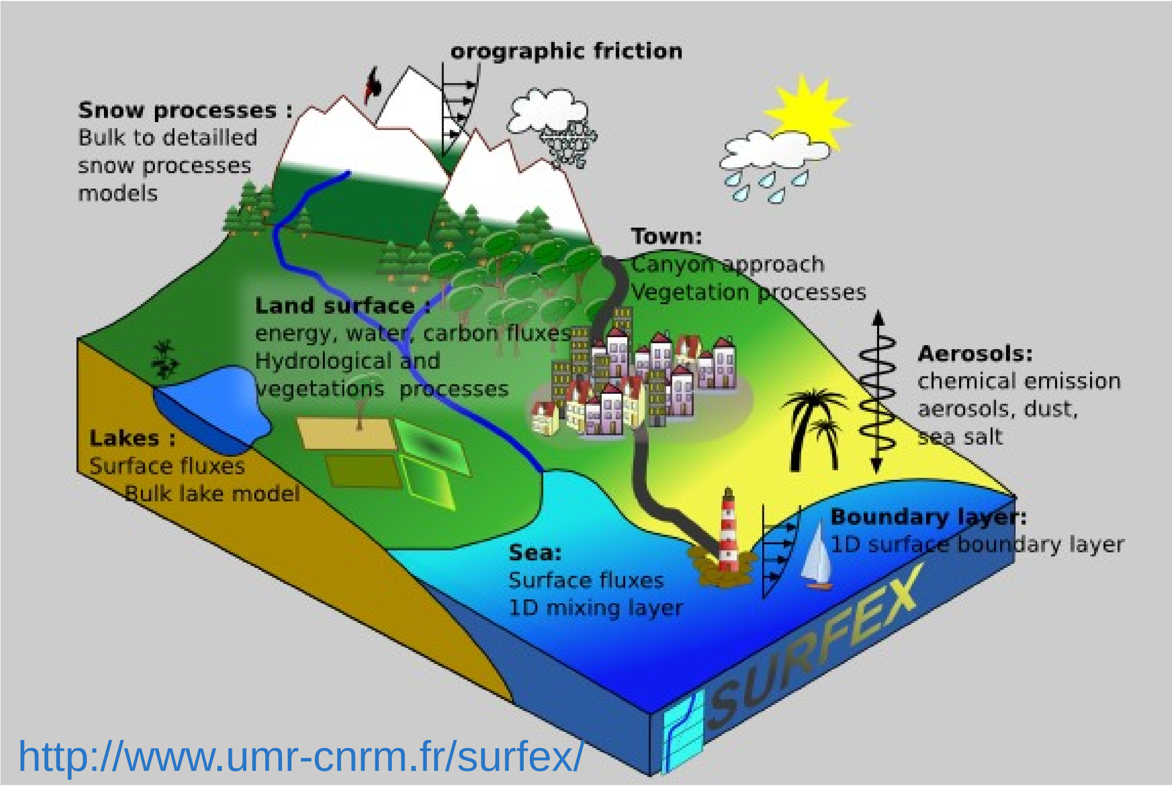


Processes by SURFEX



http://www.umr-cnrm.fr/surfex/

SURFEX/SODA and their options



SURFEX is designed to work from ESM scale to very high resolution and offline. The exact combination of options depends on application!

SURFEX physics

4 tiles:
Land, Town, Sea, Lake

Land tile:
- 1-19 patches (land use)
- 4 snow schemes
- Force-restore and diffusion soil
- Explicit canopy (MEB)
- Several hydrological options
- Phenology (LAI) and carbon

Town:
- Town Energy Balance (TEB)

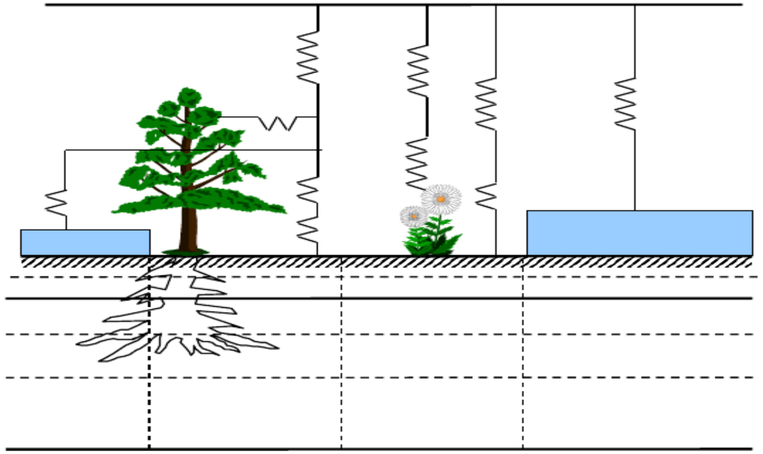
Sea:
- A few flux formulations
- Sea ice (GELATO & SICE)
- 1D ocean column model

Lake:
- FLake

Surface processes in HARMONIE-AROME (current and future) and HIRLAM

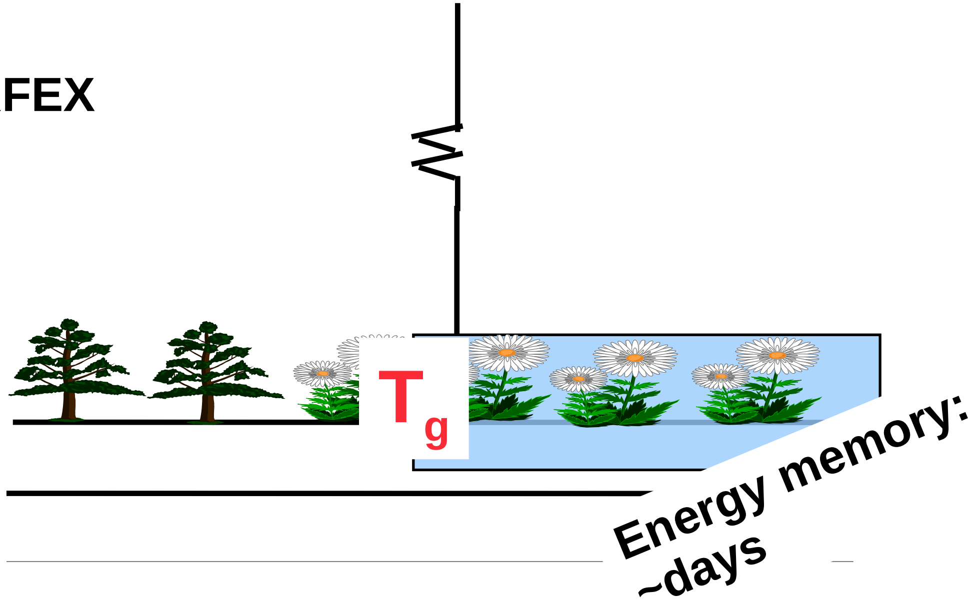
Multi level/energy

HIRLAM

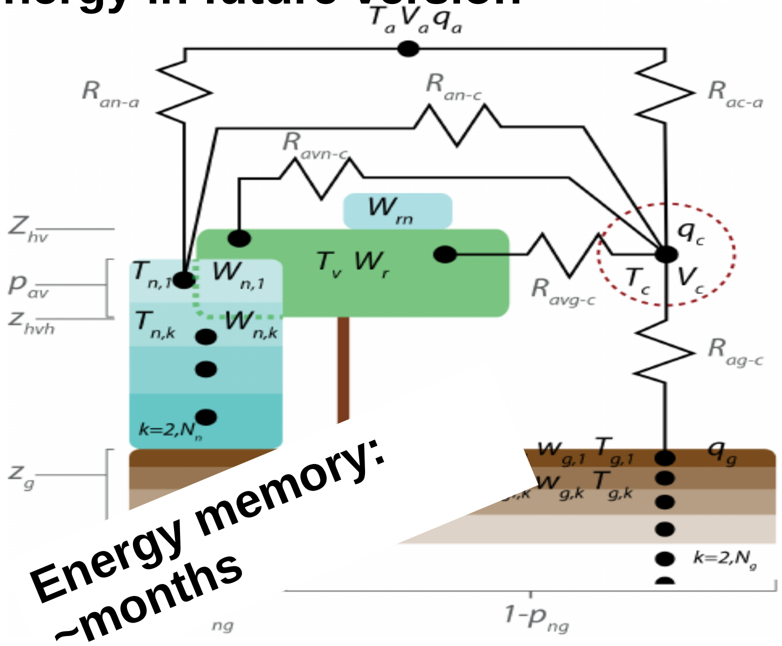


Force-restore in current operational version

SURFEX



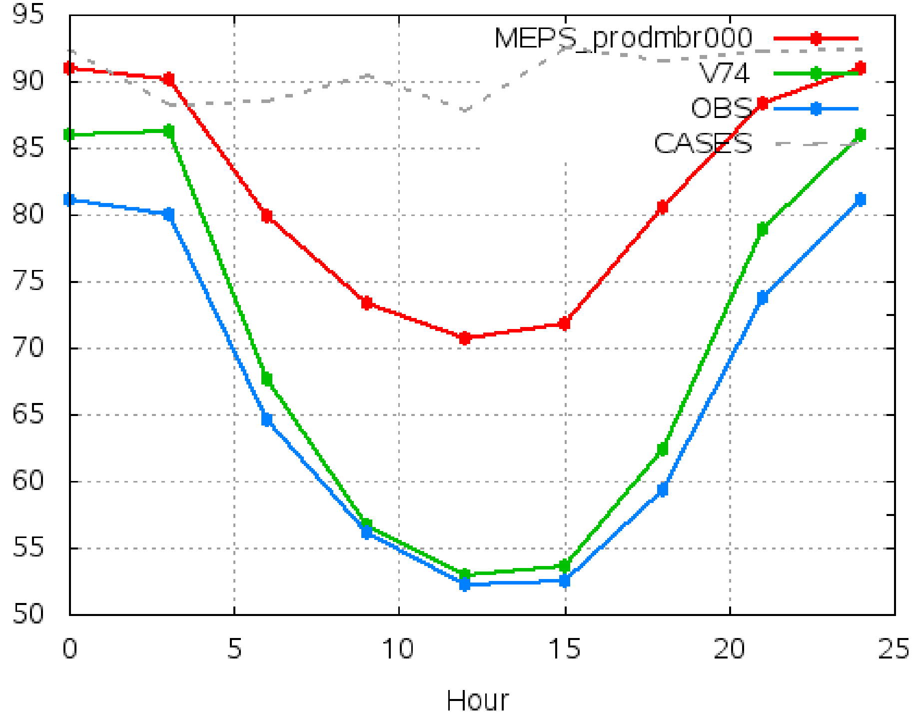
Multi level/energy in future version



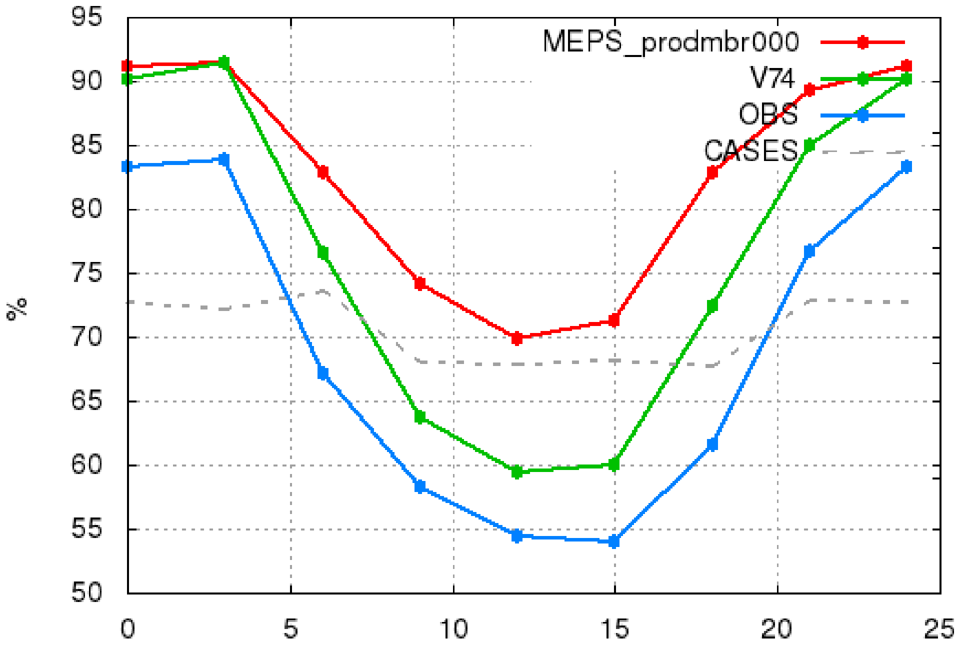
An example to motivate why we need better physics again...

Diurnal cycle of Rh2m over Finland for observations, operational HARMONIE-AROME and operational HIRLAMv7.4:

May 2017



May 2018

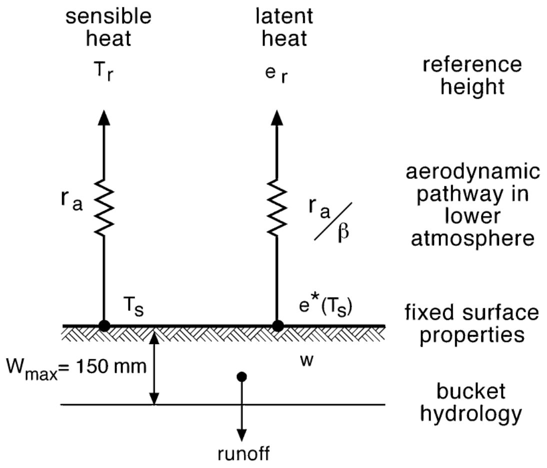


It is believed that the positive daytime bias in HARMONIE-AROME is partly due to too warm soil (due to shallow soil and short memory) which makes soil water available for transpiration. While HIRLAM still has cooler soil (longer memory) which prevents vegetation to become active.

Details in physical processes

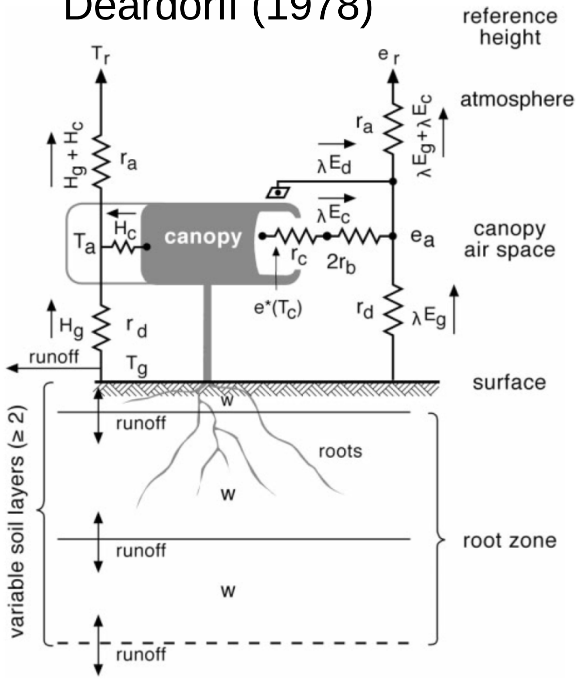
1st generation

Manabe (1969)



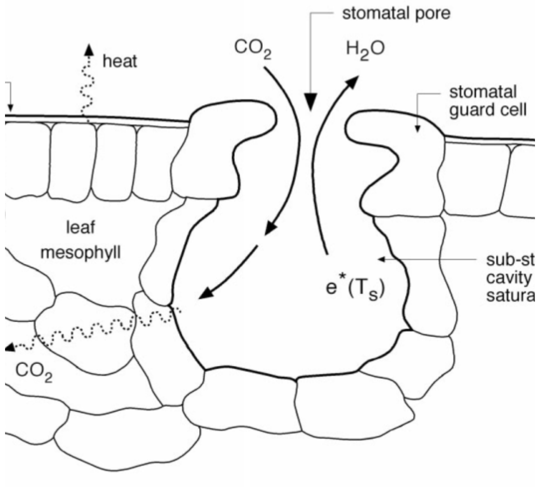
2nd generation
with explicit canopy

Deardorff (1978)



3rd generation
with carbon

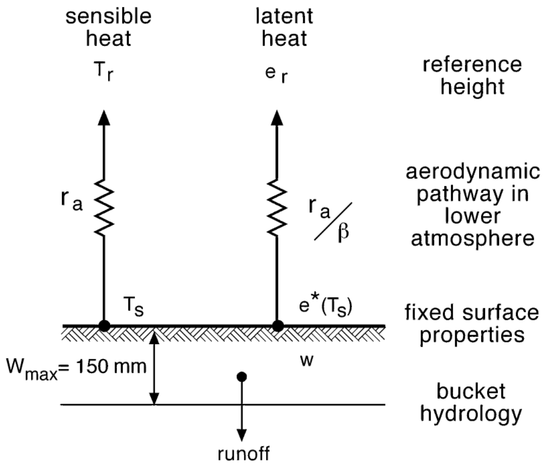
Collatz *et al*. (1991);
Sellers *et al*. (1992)



Details in physical processes

1st generation

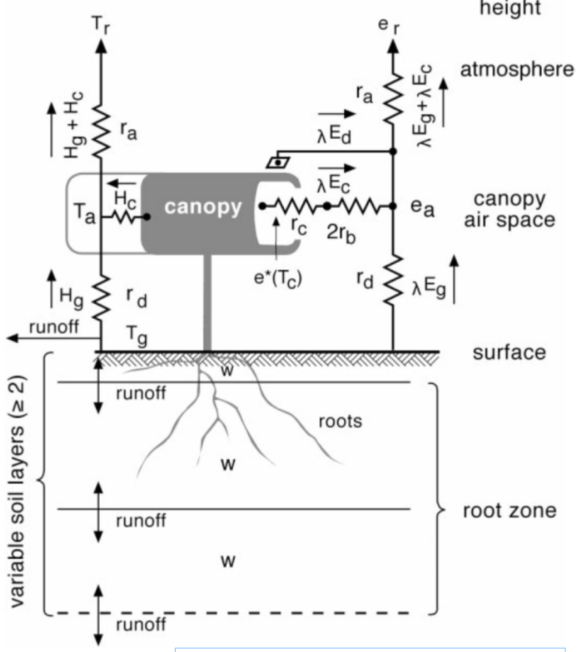
Manabe (1969)



2nd generation

with explicit canopy

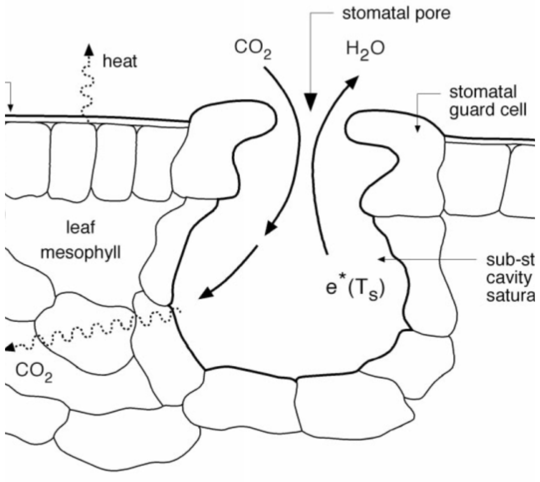
Deardorff (1978)



3rd generation

with carbon

Collatz *et al*. (1991);
Sellers *et al*. (1992)



Latest HIRLAM at 5km resolution

cy40t/h uses SURFEXv7.3 at 1.3/2.5 km resolution

ECMWF (H-TESSEL)

cy43t uses SURFEXv8 as 1.5 generation

cy43h will use SURFEXv8.1 as 2nd generation

SURFEXv8

ECMWF (C-TESSEL)

Arpege ESM uses SURFEXv8 as 3rd generation

The ALADIN-HIRLAM system: HARMONIE-AROME

The ALADIN-HIRLAM NWP system (based on IFS) includes a few configurations: ALARO, AROME-France, HARMONIE-AROME, Climate

The latest release of the HARMONIE-AROME configuration is based on cy40h.

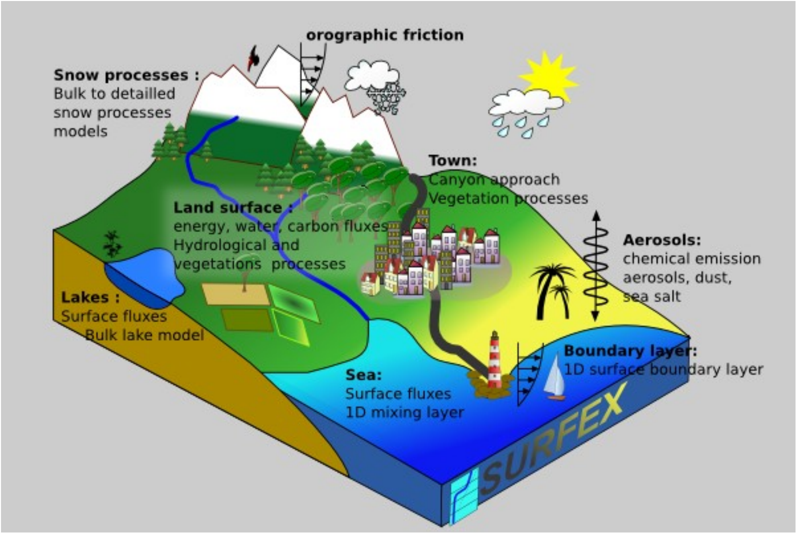
The surface perspective of the cycles:

	cy40h	**cyxxh (long term ambition)**
Land		
Patches	1 or 2	2-4 patches with explicit canopy
Soil	Force-restore	Diffusion (14 layers)
Snow	D95 (composite)	Explicit snow (12 layers)
Glacier	“Pile of snow”	Explicit snow as glacier
Assimilation	CANARI-OI	MESCAN/gridpp-EKF/EnKF
Sea	SICE	Sea ice (SICE)
Lake	FLake (optional)	FLake (later with EKF)
Town	TEB	TEB (more options)
Physiog.	ECOCLIMAP-II	ECOCLIMAP 2^{nd} generation
	1 km resolution	300 m resolution

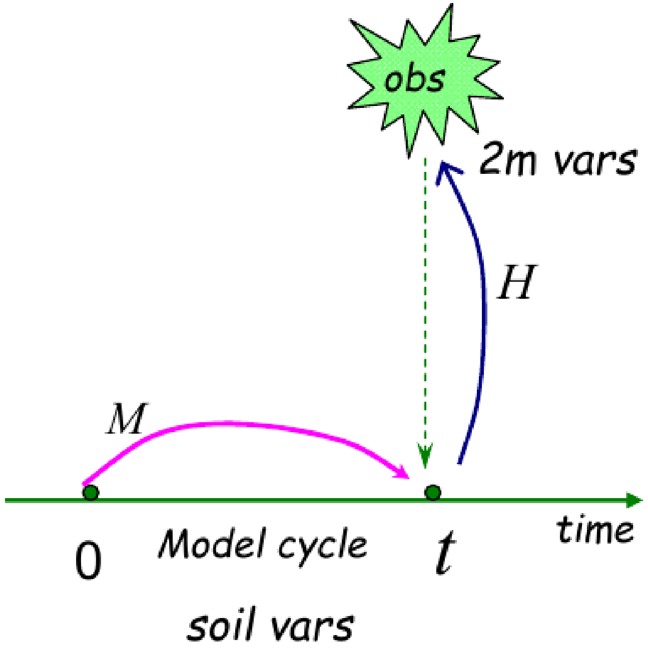
The three main components of surface modelling for NWP

Physiography

Processes



Assimilation



SURFEXv8.0 with its patches (sub-tiles) for the land tile

ECOCLIMAP 19 vegetation types

A number of parameters are associated with each vegetation type (e.g. LAI, albedo, root depth, tree height, Rsmin,...)

Bare soil
Rocks
Permanent snow/ice
C3 crops
C4 crops
Grassland
Boreal grass
Tropical grass
Irrigated crops
Wetlands, parks, irrigated grass
Temperate broadl. cold-dec. summergr
Tropical broadleaf deciduous
Temperate broadleaf evergreen
Boreal broadl. cold-dec. summergr.
Shrubs
Tropical broadleaf evergreen
Boreal needleleaf evergreen
Temperate needleleaf evergreen
Boreal needlel. cold-dec. summergr.

SURFEXv8.0 with its patches (sub-tiles) for the land tile

Rules of parameter aggregation

ECOCLIMAP 19 vegetation types

A number of parameters are associated with each vegetation type (e.g. LAI, albedo, root depth, tree height, Rsmin,...)

Number of patches specified (NPATCH)	19	12	11	10	9	8	7	6	5	4	3	2	1
Bare soil	1	1	1	1	1	1	1	1	1	1	1	2	1
Rocks	2	2	2	1	1	1	1	1	1	1	1	2	1
Permanent snow/ice	3	3	3	2	2	2	2	1	1	1	1	2	1
C3 crops	7	7	7	6	5	4	4	3	3	3	3	2	1
C4 crops	8	8	8	7	6	5	4	3	3	3	3	2	1
Grassland	10	10	10	9	8	7	6	5	5	3	3	2	1
Boreal grass	18	10	10	9	8	7	6	5	5	3	3	2	1
Tropical grass	11	11	10	9	8	7	6	5	5	3	3	2	1
Irrigated crops	9	9	9	8	7	6	5	4	4	4	3	2	1
Wetlands, parks, irrigated grass	12	12	11	10	9	8	7	6	4	4	3	2	1
Temperate broadl. cold-dec. summergr	4	4	4	3	3	3	3	2	2	2	2	1	1
Tropical broadleaf deciduous	13	4	4	3	3	3	3	2	2	2	2	1	1
Temperate broadleaf evergreen	14	4	4	3	3	3	3	2	2	2	2	1	1
Boreal broadl. cold-dec. summergr.	16	4	4	3	3	3	3	2	2	2	2	1	1
Shrubs	19	4	4	3	3	3	3	2	2	2	2	1	1
Tropical broadleaf evergreen	6	6	6	5	3	3	3	2	2	2	2	1	1
Boreal needleleaf evergreen	5	5	5	4	4	3	3	2	2	2	2	1	1
Temperate needleleaf evergreen	15	5	5	4	4	3	3	2	2	2	2	1	1
Boreal needlel. cold-dec. summergr.	17	5	5	4	4	3	3	2	2	2	2	1	1

SURFEXv8.0 with its patches (sub-tiles) for the land tile

Rules of parameter aggregation

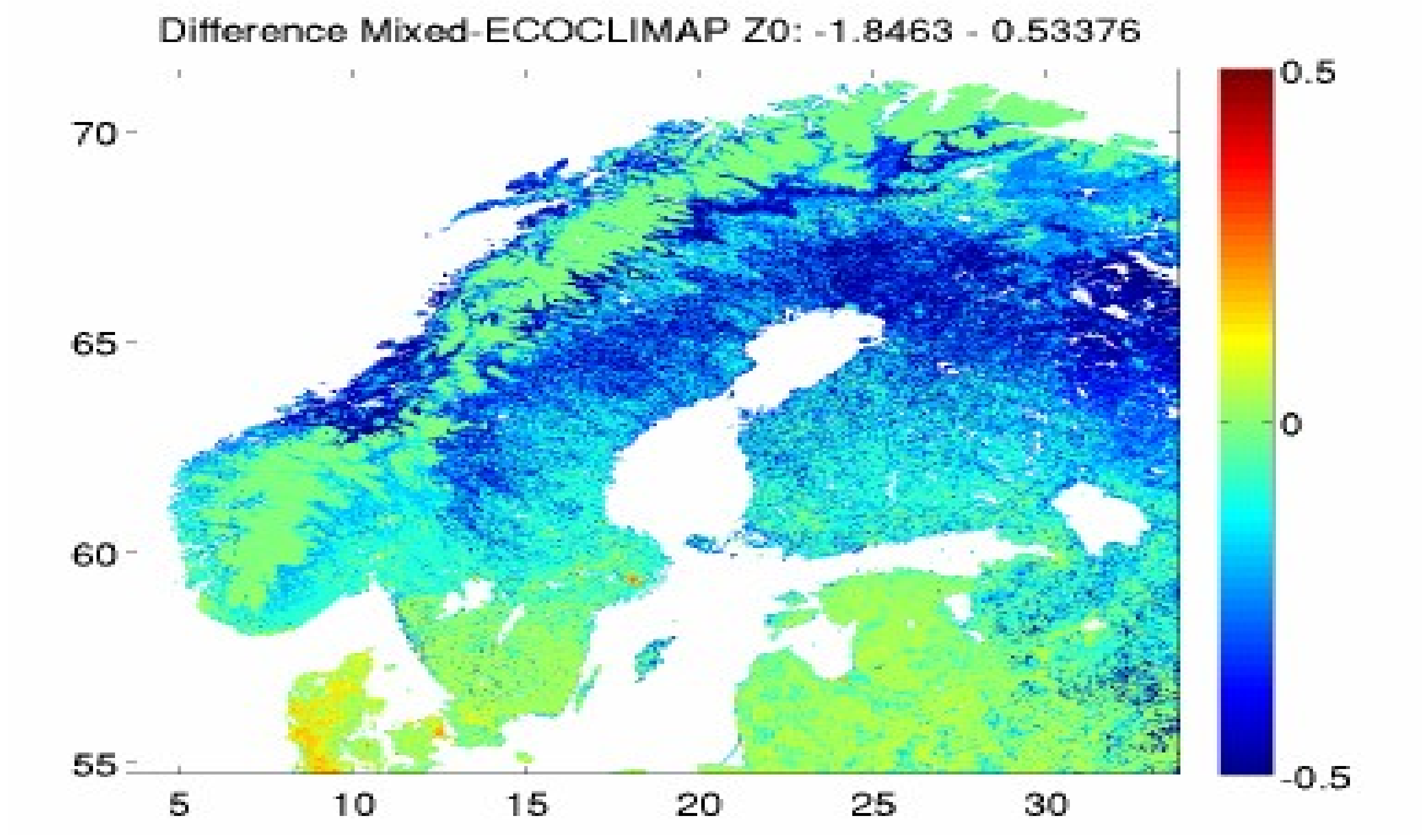
A number of parameters are associated with each vegetation type (e.g. LAI, albedo, root depth, tree height, Rsmin,...)

ECOCLIMAP 19 vegetation types		Arpege climate(?)										MetCoOp	HARMONIE-AROME default
Number of patches specified (NPATCH)	19	12	11	10	9	8	7	6	5	4	3	2	1
Bare soil	1	1	1	1	1	1	1	1	1	1	1	2	1
Rocks	2	2	2	1	1	1	1	1	1	1	1	2	1
Permanent snow/ice	3	3	3	2	2	2	2	1	1	1	1	2	1
C3 crops	7	7	7	6	5	4	4	3	3	3	3	2	1
C4 crops	8	8	8	7	6	5	4	3	3	3	3	2	1
Grassland	10	10	10	9	8	7	6	5	5	3	3	2	1
Boreal grass	18	10	10	9	8	7	6	5	5	3	3	2	1
Tropical grass	11	11	10	9	8	7	6	5	5	3	3	2	1
Irrigated crops	9	9	9	8	7	6	5	4	4	4	3	2	1
Wetlands, parks, irrigated grass	12	12	11	10	9	8	7	6	4	4	3	2	1
Temperate broadl. cold-dec. summergr	4	4	4	3	3	3	3	2	2	2	2	1	1
Tropical broadleaf deciduous	13	4	4	3	3	3	3	2	2	2	2	1	1
Temperate broadleaf evergreen	14	4	4	3	3	3	3	2	2	2	2	1	1
Boreal broadl. cold-dec. summergr.	16	4	4	3	3	3	3	2	2	2	2	1	1
Shrubs	19	4	4	3	3	3	3	2	2	2	2	1	1
Tropical broadleaf evergreen	6	6	6	5	3	3	3	2	2	2	2	1	1
Boreal needleleaf evergreen	5	5	5	4	4	3	3	2	2	2	2	1	1
Temperate needleleaf evergreen	15	5	5	4	4	3	3	2	2	2	2	1	1
Boreal needlel. cold-dec. summergr.	17	5	5	4	4	3	3	2	2	2	2	1	1

However, parameter values can be dubious...

During a development period we concluded that the ECOCLIMAP tree height (roughness) seemed to be overestimated in the northern half of the MetCoOp domain. Therefore, ECOCLIMAP tree height was replaced by laser estimated tree height. See article in No 10 of ALADIN/HIRLAM Newsletter: www.umr-cnrm.fr/aladin/IMG/pdf/nl10.pdf.

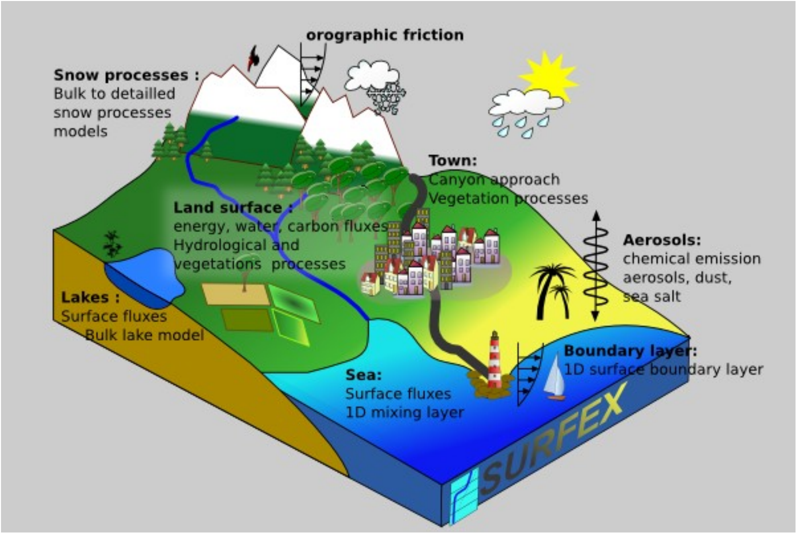
Difference (new – old) roughness :



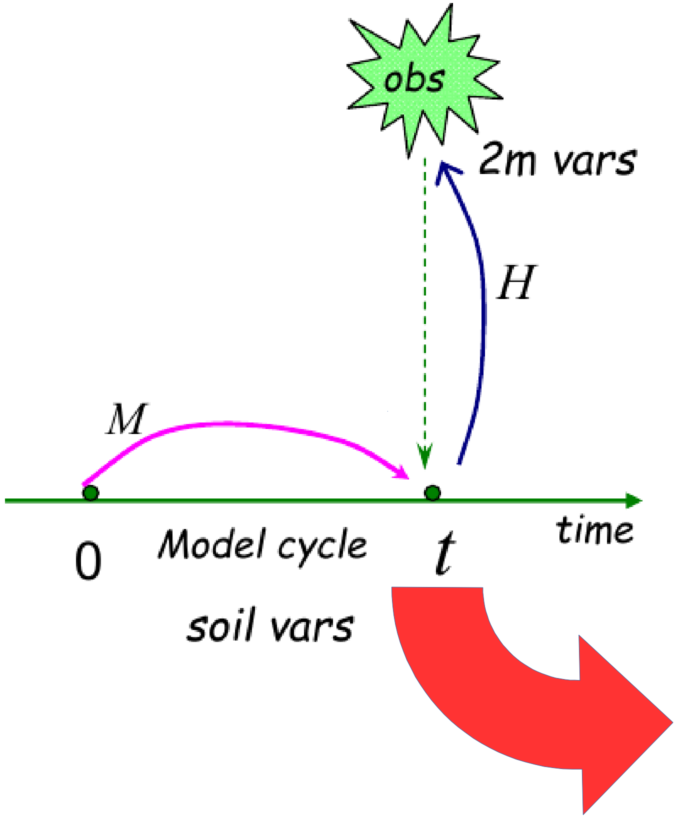
The three main components of surface modelling for NWP

Physiography

Processes



Assimilation



Surface data assimilation – Principles

After a few hours of integration (here 3 hours) the prognostic variables in the soil have deviated from the truth (which we don't know) due to non-linearities of the system but also due to biases in the system.

$$\mathbf{X}_{\mathbf{A}} = \mathbf{X}_{\mathbf{B}} + \mathbf{K}\left(\mathbf{Y} - H\left(\mathbf{X}_{\mathbf{B}}\right)\right)$$

The job of the surface data assimilation is to utilize observations (currently T2m and Rh2m) to push a subset of the prognostic variables (the control variables) towards a new analysed state.

Surface data assimilation – Principles

After a few hours of integration (here 3 hours) the prognostic variables in the soil have deviated from the truth (which we don't know) due to non-linearities of the system but also due to biases in the system.

$$\mathbf{X_A} = \mathbf{X_B} + \mathbf{K}(\mathbf{Y} - H(\mathbf{X_B}))$$

The job of the surface data assimilation is to utilize observations (currently T2m and Rh2m) to push a subset of the prognostic variables (the control variables) towards a new analysed state.

Here, *H*, in $H(\mathbf{X_B})$ represents the forward model or observation operator. I.e., the translation from model variables to observations if the variables representing the observations do not exist in the model already.

Surface data assimilation – Principles

After a few hours of integration (here 3 hours) the prognostic variables in the soil have deviated from the truth (which we don't know) due to non-linearities of the system but also due to biases in the system.

$$\mathbf{X}_{\mathbf{A}} = \mathbf{X}_{\mathbf{B}} + \mathbf{K}\left(\mathbf{Y} - H\left(\mathbf{X}_{\mathbf{B}}\right)\right)$$

The job of the surface data assimilation is to utilize observations (currently T2m and Rh2m) to push a subset of the prognostic variables (the control variables) towards a new analysed state.

If the observation errors are big compared to the model errors the new analysed state will be close to the first guess (or background, the value of the control variables after 3 hours in the previous forecast). On the other hand, if the model errors are big compared to the observation errors the new analysed state will be close to the observations.

Here, H, in $H\left(\mathbf{X}_{\mathbf{B}}\right)$ represents the forward model or observation operator. I.e., the translation from model variables to observations if the variables representing the observations do not exist in the model already.

Surface data assimilation – Principles

After a few hours of integration (here 3 hours) the prognostic variables in the soil have deviated from the truth (which we don't know) due to non-linearities of the system but also due to biases in the system.

$$\mathbf{X}_A = \mathbf{X}_B + \mathbf{K}\left(\mathbf{Y} - H\left(\mathbf{X}_B\right)\right)$$

The job of the surface data assimilation is to utilize observations (currently T2m and Rh2m) to push a subset of the prognostic variables (the control variables) towards a new analysed state.

If the observation errors are big compared to the model errors the new analysed state will be close to the first guess (or background, the value of the control variables after 3 hours in the previous forecast). On the other hand, if the model errors are big compared to the observation errors the new analysed state will be close to the observations.

Here, *H*, in $H\left(\mathbf{X}_B\right)$ represents the forward model or observation operator. I.e., the translation from model variables to observations if the variables representing the observations do not exist in the model already.

It is the Kalman Gain, K, which represents the relationship between observation errors and model errors. E.g., a small K (small/big model/observation error) gives higher weight to the first guess.

Surface data assimilation – Optimal Interpolation (OI)

In the current operational HARMONIE-AROME we use Optimal Interpolation (OI) as our surface data assimilation method. Here the Kalman Gain, K, is simply calibrated to find a relationship between observations and control variables.

Observations used are:

- **T2m**
- **Rh2m**

The control variables are, i.e. the prognostic variables which are assimilated:

- **TG1 (surface temperature)**
- **TG2 (deep soil temperature)**
- **WG1 (surface soil moisture)**
- **WG2 (deep soil moisture)**

Surface data assimilation – Optimal Interpolation (OI)

In the current operational HARMONIE-AROME we use Optimal Interpolation (OI) as our surface data assimilation method. Here the Kalman Gain, K, is simply calibrated to find a relationship between observations and control variables.

Observations used are:

- **T2m**
- **Rh2m**

The control variables are, i.e. the prognostic variables which are assimilated:

- **TG1 (surface temperature)**
- **TG2 (deep soil temperature)**
- **WG1 (surface soil moisture)**
- **WG2 (deep soil moisture)**

Such a calibration is difficult to make general and good for all possible situations (dry/moist soil, snow or no snow conditions, bare soil / inactive vegetation / active vegetation, …)

Also, OI is difficult to expand to many control variables and to a combination of different types of observations (e.g. SYNOP and satellite).

Surface data assimilation – Simplified Extended Kalman Filter (SEKF)

Next step for HARMONIE-AROME is to apply Simplified Extended Kalman Filter (SEKF). Here the Kalman Gain, K, is calculated for each assimilation cycle to find the current relationship between observations and control variables:

$$\mathbf{K} = \mathbf{B} \cdot \mathbf{H}^{\mathsf{T}} \left(\mathbf{H} \cdot \mathbf{B} \cdot \mathbf{H}^{\mathsf{T}} + \mathbf{R} \right)^{-1}$$

Here, B represents the model (background) errors and R the observation errors. H is the Jacobian, which is estimated by running SURFEX offline for as many control variables that exist, where for each run, one control variable at a time is perturbed. Gives e.g. dT2m/dTG2.

Surface data assimilation – Simplified Extended Kalman Filter (SEKF)

Next step for HARMONIE-AROME is to apply Simplified Extended Kalman Filter (SEKF). Here the Kalman Gain, K, is calculated for each assimilation cycle to find the current relationship between observations and control variables:

$$\mathbf{K} = \mathbf{B} \cdot \mathbf{H}^{T} \left(\mathbf{H} \cdot \mathbf{B} \cdot \mathbf{H}^{T} + \mathbf{R} \right)^{-1}$$

Here, B represents the model (background) errors and R the observation errors. H is the Jacobian, which is estimated by running SURFEX offline for as many control variables that exist, where for each run, one control variable at a time is perturbed. Gives e.g. dT2m/dTG2.

Pros:

- **the current situation is used to estimate K. No calibration needed.**
- **New processes can easily be connected to the assimilation.**
- **Allows for several types of observations at the same time.**

Cons:

- **If the system behaves strongly non-linear, e.g. during precipitation events, the calculation of the Jacobians can break down. Safety actions are needed.**
- **Can become expensive if too many control variables are used.**

Surface data assimilation – Satellite observations

ASCAT

We can utilize satellite data in two ways:

- **Use a product where someone else has processed the satellite data to estimate e.g. soil moisture or snow extent. Not too complicated but another model is usually in between which increase uncertainties. Ongoing examples are:**

Instrument	Satellite	Product	Comment
ASCAT	METOP-A/B	Surface soil moisture	
SEVIRI	MSG	Snow extent (H-SAF)	
SMMR, SSM/I, AVHRR GAC	several	Snow extent	Not for operational use

Surface data assimilation – Satellite observations

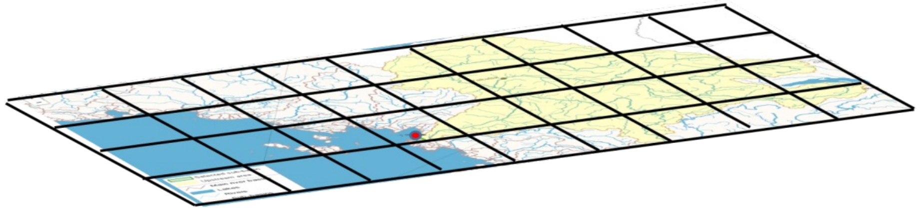
MIRAS/
SMOS

We can utilize satellite data in two ways:

- **Use radiances directly. Good, less dependent on other models but requires a forward model which translate the model state to satellite radiances. Test and development are ongoing for:**

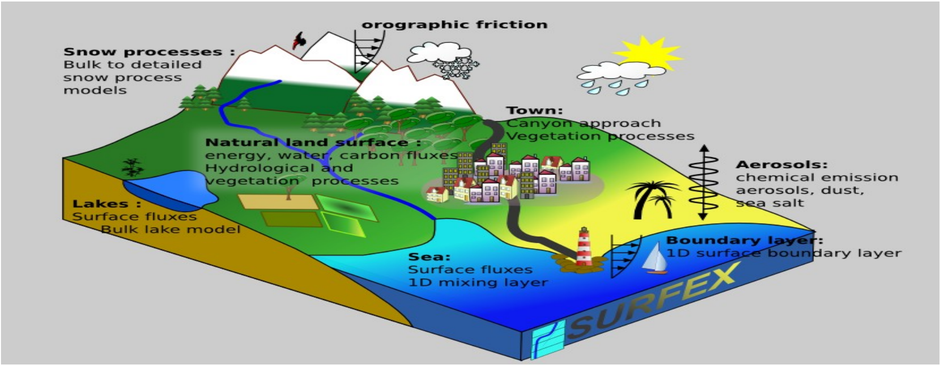
Instrument	**Satellite**	**Intention**	**Forward model *H*(X_b)**
AMSR-2	GCOM-W1	surface soil moisture (7 GHz) shallow snow (89 GHz) moderate snow (37 GHz) deep snow (10, 19 GHz)	FASTEM + RTTOV?
MIRAS	SMOS	surface soil moisture (1.4 GHz)	CMEM + FASTEM (water)?
SAR-C (Sentinel-1)	Sentinel-1A	wet snow snow extent	MEMLS3-A (Proksch et al. 2015)

Nordic hydrometeorological system



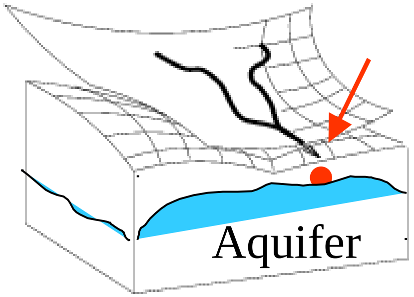
Coupled to NWP/climate model or Offline with interpolation/downscaling to higher resolution and/or spatial/time correction or boundary layer processes...

MESAN analysis system + STRÅNG radiation analysis, MetCoOp operational + EPS, HARMONIE-Climate, ...



SURFEX Nature-Town-Lake-Sea Nature potential: 19 patches, multi-layer snow and soil, explicit veg/snow, veg dynamics

OASIS coupler for deep and surface runoff, ground water, flooding, irrigation, ...



Hydrological ground water and routing model, e.g. HYPE ”light”, CaMaFlood, something new,…

Coupled Ensemble data assimilation system

Puffin at Cape Ingólfshöfði, southern Iceland
June 12th 2018

Takk!

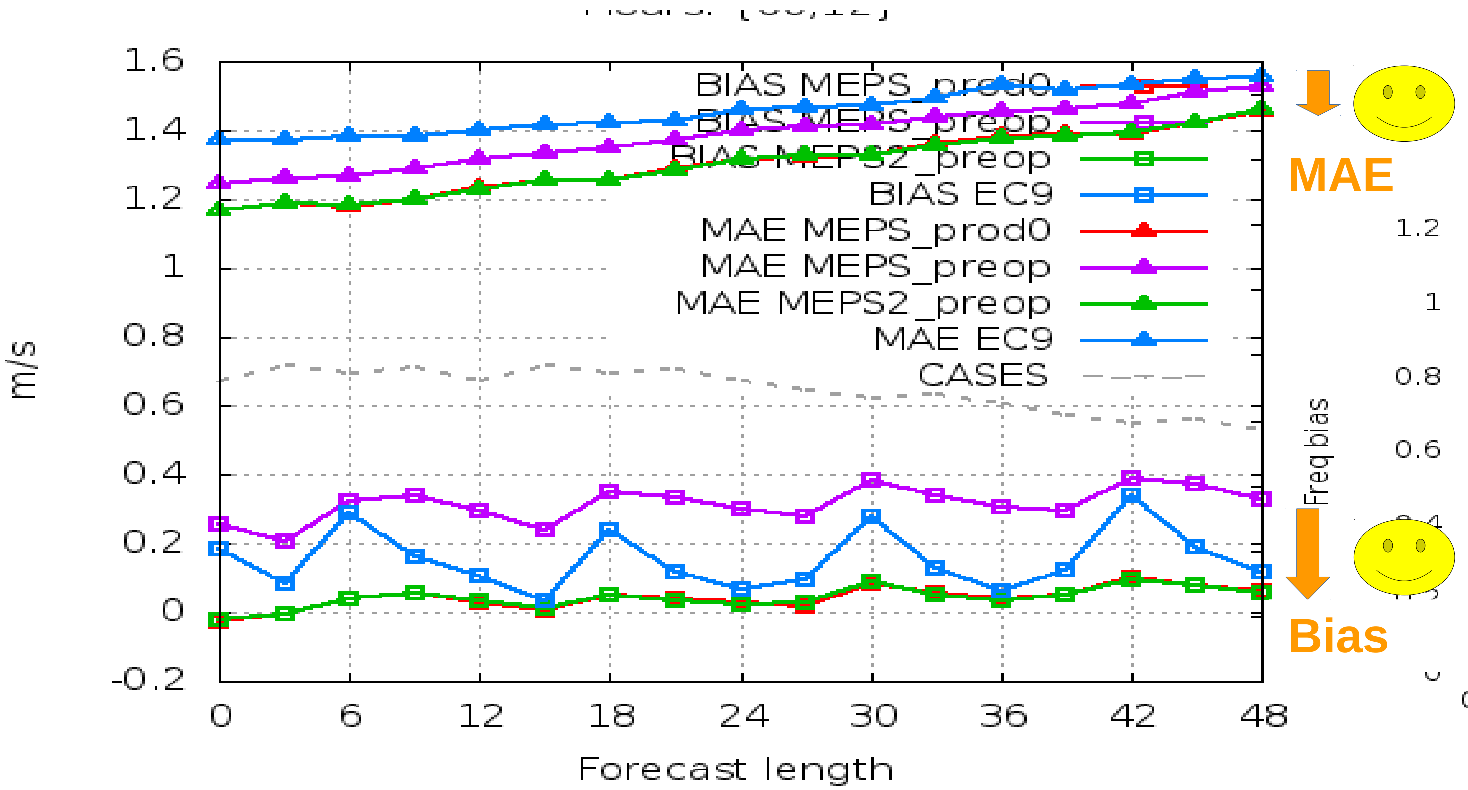
Kiitos!

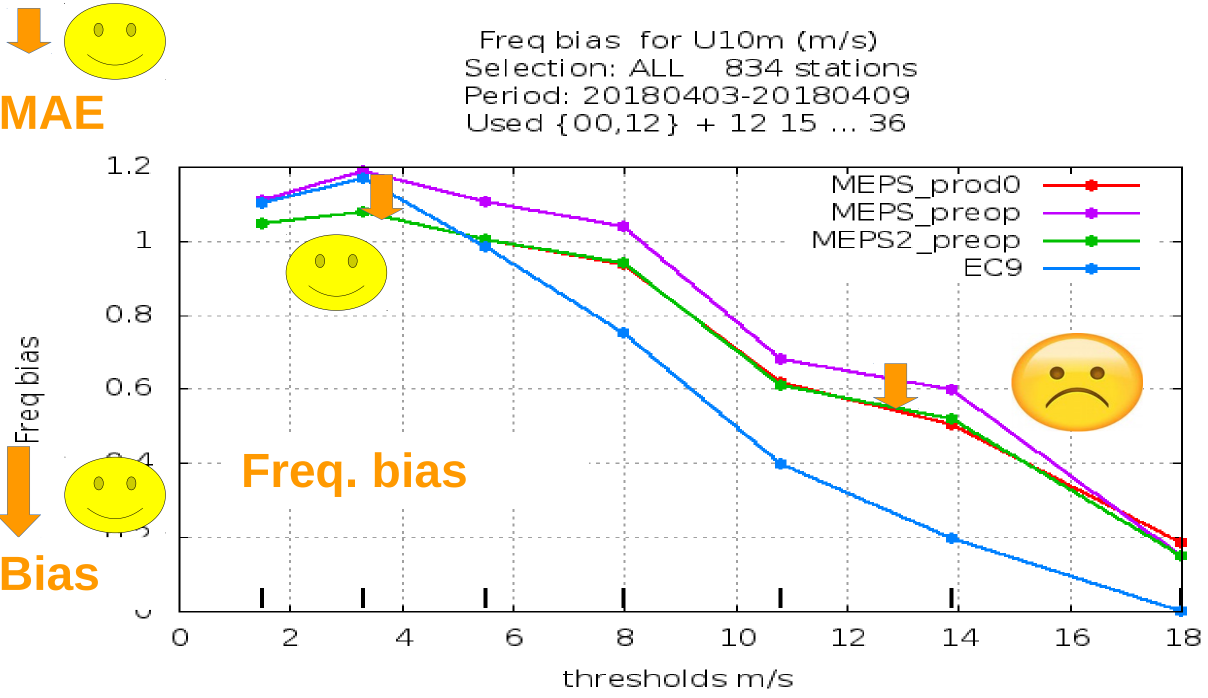
Tak!

Tack!

With modified tree height/roughness U10m improves

U10m: 7 days statistics from last week over the MetCoOp domain





Operational (cy40h1.1.1): No SBL, 2 patches, FLake
Previous version: SBL, 1 patch, TG2 for lake water
ECMWF

People involved: Patrick Samuelsson, Mariken Homleid, Trygve Aspelien, Ulf Andrae, Matti Horttanainen (FMI).