https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

A 10 km North American Precipitation and Land Surface Reanalysis Based on the GEM Atmospheric Model

Nicolas Gasset[1,*], Vincent Fortin[1,*], Milena Dimitrijevic[1], Marco Carrera[1], Bernard Bilodeau[1], Ryan Muncaster[1], Étienne Gaborit[1], Guy Roy[2], Nedka Pentcheva[2], Maxim Bulat[2], Xihong Wang[2], Radenko Pavlovic[2], Franck Lespinas[2], and Dikra Khedhaouiria[1]

[*]These authors contributed equally to this work.
[1]Meteorological Research Division, Environment and Climate Change Canada, Dorval, QC, Canada
[2]Meteorological Service of Canada, Environment and Climate Change Canada, Dorval, QC, Canada

Correspondence: Vincent Fortin (vincent.fortin@canada.ca)

Abstract. Environment and Climate Change Canada has initiated the production of a 1980-2018, 10 km, North American precipitation and surface reanalysis. ERA-Interim is used to initialize the Global Deterministic Reforecast System (GDRS) at a 39 km resolution. Its output is then dynamically downscaled to 10 km by the Regional Deterministic Reforecast System (RDRS). Coupled with the RDRS, the Canadian Land Data Assimilation System (CaLDAS) and Precipitation Analysis (CaPA) are used to produce surface and precipitation analyses. All systems used are close to operational model versions and configurations. In this study, a 7-year sample of the reanalysis (2011-2017) is evaluated. Verification results show that the skill of the RDRS is stable over time, and equivalent to that of the current operational system. The impact of the coupling between RDRS and CaLDAS is explored using an early version of the reanalysis system which was run at 15 km resolution for the period 2010-2014, with and without the use of CaLDAS. Significant improvements are observed with CaLDAS in the lower troposphere and surface layer, especially for the 850 hPa dew point and absolute temperatures in summer. Precipitation is further improved through an offline precipitation analysis which allows the assimilation of additional observations of 24-h precipitation totals. The final dataset should be of particular interest for hydrological applications focusing on trans-boundary and northern watersheds, where existing products often show discontinuities at the border and assimilate very few - if any - precipitation observations.

1 Introduction

Atmospheric reanalysis datasets are invaluable tools allowing for better understanding and analysis of global and regional water cycles by integrating data assimilation techniques with state-of-the-art numerical models of the atmosphere and Earth's surface observations of the water and energy cycle. Indeed, this was one of the main objectives in the design and the development of recent reanalysis datasets such as NASA's Modern-Era Retrospective analysis for Research and Applications (MERRA) (Bosilovich et al., 2008; Rienecker et al., 2011; Bosilovich et al., 2011; Takacs et al., 2016), ECMWF's ERA-Interim reanalysis (Dee et al., 2011) and ERA5 (Hersbach et al., 2020).

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Data assimilation is the process by which a system state estimated by a model is merged with observations of that system, in order to adjust the trajectory of the model (Fletcher, 2017). More specifically, a gridded numerical model output (commonly referred to as the background or trial field) is combined with observations of all kinds, through advanced mathematical methods constrained by the laws of physics in order to obtain a gridded representation (generally referred to as an analysis) as close as possible to the true state which the numerical model intends to predict and observations intend to measure.

Atmospheric reanalysis datasets are generally obtained by performing data assimilation of atmospheric observations every few hours and integrating a numerical weather prediction model (NWP) between analysis times. The NWP model itself relies on a land surface scheme to represent energy, mass and momentum exchanges occurring at the Earth's surface. Most reanalyses are designed to provide optimal results for large scale atmospheric fields, sometimes at the price of degradation of surface prediction skill and bias (Albergel et al., 2013; Fujiwara et al., 2017). In addition, reanalyses often do not assimilate any observations of precipitation and of the land-surface state, but instead only provide short-term forecasts of these variables.

Another limitation of existing atmospheric reanalyses is that their spatial resolution can be too coarse to be fully suited for land-surface applications at the regional scale (Marke et al., 2011), even for the most recent and advanced ones such as ERA5 (Hersbach et al., 2020) or the North American Regional Reanalysis (NARR; Mesinger et al., 2006). In order to obtain predictions at the scale required for many land-surface and hydrological modelling applications, higher resolution atmospheric forcing and land-surface model outputs are necessary.

As a consequence, separate/complementary land surface reanalyses have been developed by both NASA and ECMWF where the above mentioned atmospheric reanalyses are directly used to force advanced land surface models in an open-loop manner and at higher horizontal resolution. Considering that precipitation is one of the most influential forcing for land-surface processes, precipitation is generally adjusted and locally rescaled based on observation datasets such as the Global Precipitation Climatology Project pentad – GPCP v2.1 (Adler et al., 2003). Such datasets, including MERRA-Land (Reichle et al., 2011), ERA-Interim-Land (Balsamo et al., 2015) and ERA5-Land (Muñoz Sabater et al., 2018) are shown to improve surface prediction when compared to the underlying original reanalysis. However, with the exception of a simple precipitation adjustment, data assimilation of neither surface nor remotely sensed observations is involved in the process.

The main objectives of this paper are (1) to present a strategy for the production, at relatively low computing cost, of high-resolution surface and precipitation regional reanalyses which includes two-way coupling between the land data assimilation system and the NWP model as well as assimilation of precipitation observations, and (2) to evaluate a North-American reanalysis based on the Global Environmental Multiscale (GEM model, Côté et al., 1998b, a; Girard et al., 2014) model obtained using the proposed methodology, and currently under production at the Canadian Centre for Meteorological and Environmental Prediction (CCMEP) of Environment and Climate Change Canada (ECCC).

The desired outputs from this system are hourly time series, available on a ~ 10 km regular grid of the main meteorological variables that are required in order to operate surface and hydrology models: near-surface air temperature, near-surface air humidity, surface pressure, wind speed and direction, incoming infrared flux at the surface, downward solar flux at the surface, and finally precipitation.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

One of the important motivations for the production of this reanalysis is, for ECCC, to better support the International Joint Commission (IJC) which prevents and resolves disputes related to transboundary watersheds covered by the Canada-US Boundary Waters Treaty. Too often, North American climate datasets show discontinuities at the Canada-US border (Gronewold et al., 2017). By covering seamlessly all of North America, this product is expected to facilitate hydrometeorological and hydroclimatic studies involving transboundary watersheds in North America.

1.1 Open-loop land-surface simulations as an alternative to reanalysis land-surface data

The approach used by ERA5-Land, MERRA-Land and ERA-Interim-Land of running a land-surface model in open-loop with adapted meteorological forcing data is fairly common in the land surface and hydrology modelling community (Albergel et al., 2013; Muñoz Sabater et al., 2018), i.e. surface models are often integrated in an open-loop manner forced by some kind of atmospheric gridded datasets in order to predict the state of the land surface. There are several major advantages to this approach: (1) atmospheric forcing can be improved through post-processing, (2) horizontal resolution can be increased without a prohibitive computational cost, and (3) a more recent land-surface scheme, or perhaps more adapted to the problem at hand, can be used. While convenient and relatively straightforward, results from this approach are fully dependent on both the forcing data quality and the abilities of the land surface model when used with this specific forcing dataset to reproduce the processes of interest. No observations are directly involved other than those used to produce the atmospheric forcing. Thus the land surface model errors can grow over time since the land-surface model is allowed to drift without the constraints from surface observations. These errors can be caused by limitations of the land-surface scheme but also by an imbalance between the atmospheric forcing obtained from a reanalysis product and the surface feedback predicted by the land-surface model when a different horizontal resolution or different model is used.

1.2 Dynamical downscaling of reanalysis data to improve horizontal resolution

The resolution limitation of existing reanalysis products can be overcome through dynamical downscaling, using either a limited-area NWP model or a regional climate model. This is once again fairly common in the literature, see for example Yoshimura and Kanamitsu (2008) and Giorgi (2019). Such models, when initialized and/or forced at their boundaries with reanalysis data have been shown to perform properly and can improve results (Lucas-Picher et al., 2013; Chikhar and Gauthier, 2014). When the objective is to generate high-resolution datasets covering periods of more than a few years, this approach is generally preferable to the use of archived outputs from operational analysis and/or forecasting systems. Indeed, although operational analyses and forecasts can be more skillful than reanalyses, these datasets are not consistent in time, i.e. their time/space resolution and quality evolve as a function of operational implementations, preventing their usage for long term applications where consistency is required (Smith et al., 2014).

The atmospheric model used for dynamical downscaling generally relies on a land-surface model sufficiently different from the one used as part of the lower resolution reanalysis system. As a result, it cannot be initialized from land-surface state variables provided with the reanalysis. The usual solution to this problem is to continuously integrate the land-surface model, cycling the land-surface state variables after the end of each atmospheric model run. Again, the land-surface model is allowed to

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

drift with time. Furthermore, because the land-surface model interacts with the atmospheric model in a dynamical downscaling experiment, the atmospheric model itself can drift with time. It will be shown that this feature of the standard approach to dynamical downscaling of reanalyses can lead to significant atmospheric model errors and biases.

1.3 Avoiding land-surface and atmospheric model drift through land-surface data assimilation

Land-surface model drift can be controlled through the use of a land-surface data assimilation system (Carrera et al., 2015). At the same time, the surface analyses provided by such a system can be used to better initialize an atmospheric model used for dynamical downscaling, if the same land-surface model is used by the land-surface data assimilation system and the atmospheric prediction system. The dynamically downscaled atmospheric forcing can furthermore be used as input to the land-surface data assimilation scheme, thus ensuring consistency between the surface predictions performed as part of the assimilation and forecast cycles. The computing cost of such an approach is higher than that of running an open-loop simulation, but still much less than that of a complete atmospheric and land-surface reanalysis (Fairbairn et al., 2017). It will be shown that coupling a land-surface data assimilation scheme with a limited area NWP model for dynamical downscaling is well worth the effort as it leads to significantly improved near-surface atmospheric and land-surface predictions.

1.4 Structure of the paper

The present study documents a methodology for increasing the resolution of a reanalysis product through dynamical downscaling while avoiding model drift near the surface through land-surface data assimilation. Two versions of this reanalysis are introduced and evaluated: a preliminary version with a horizontal resolution of 15 km is used to assess the impact of the land-data assimilation system over the period 2010-2014, and a second version with a horizontal resolution of 10 km is compared over the period 2010-2017 to the operational NWP system which was in operation at CCMEP over the same time period. This atmospheric reforecast and land-surface reanalysis covers the whole North America and the Arctic Ocean, and the evaluation is performed over Mexico, the United States of America and Canada. In section 2, the methodology followed to obtain such a product is presented, along with the observational datasets involved in land-surface and precipitation data assimilation. Verification metrics used throughout the paper are also introduced. In section 3, the added value of surface data and precipitation assimilation versus more standard dynamical downscaling approaches is assessed both from a surface perspective and with regards to the free atmosphere, based on the 15 km configuration of the reanalysis. Section 4 focuses on the evaluation of precipitation, temperature, dew point from the 15 km and 10 km configurations, relying on surface and streamflow observations. A discussion and conclusion follow.

2 Data and Method

When designing a reanalysis system, many decisions have a significant impact on the resources and time required to perform the reanalysis. Despite these constraints, horizontal resolution must be sufficient to allow for an acceptable representation of the hydrological cycle. Benedict et al. (2019) suggest that 25 km is likely insufficient to represent the climatic driver of the

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Mississippi river basin, but Deacu et al. (2012) obtained satisfactory results when simulating the net basin supplies of the Great Lakes basin at 15 km resolution.

Based on the results of Deacu et al. (2012), initial tests aimed at fine-tuning the configuration of the reanalysis system were performed at 15 km resolution. However, a resolution of 10 km was chosen for the production of a 1980-2018 reanalysis, in order to match the current resolution of the Regional Deterministic Prediction System (RDPS) and of the Regional Ensemble Prediction System (REPS) currently in operation at CCMEP for short-term weather forecasting over North-America.

Having settled on a system operating at the meso-gamma scale over North America, the system was designed in order to provide all atmospheric forcing required to perform land-surface and hydrological modelling at the regional scale, and to be coherent with in-situ surface observations (e.g. precipitation, absolute and dew point temperatures and snow depth), while avoiding the need for atmospheric data assimilation.

As illustrated in the left panel of Fig. 1, the global reanalysis ERA-Interim (Dee et al., 2011) is used as initial atmospheric conditions of the so-called Global Deterministic Reforecast System (GDRS) in order to produce a global reforecast at higher resolution than ERA-Interim. Then, following a dynamical downscaling approach, the so-called Regional Deterministic Reforecast System (RDRS) is applied, also initialized by ERA-Interim (Dee et al., 2011) but driven by the GDRS, to produce a reforecast at higher resolution covering the whole North America and Arctic Ocean. Both the GDRS and RDRS are based on the GEM model.

ERA-Interim, and not ERA5, is used for initializing the GDRS and RDRS, mainly because ERA5 was not available during the development phase of this reanalysis project.

[Figure 1 about here.]

Surface initial conditions (sea surface temperature, sea ice concentration/thickness, soil moisture, soil temperature and snowpack conditions) consistent with the driving data and the surface scheme of the atmospheric model are also required. For the GDRS, initial land-surface conditions are obtained from an a priori offline (open-loop) of GEM's land-surface model, GEM-Surf (Carrera et al., 2010; Bernier and Bélair, 2012), directly forced by the near-surface fields of the ERA-Interim reanalysis as well as the 3-hour precipitation amounts (Gagnon and Co-authors, 2015). This offline system relies on a version of the ISBA land-surface scheme (Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996) adapted for use in the GEM model (Bélair et al., 2003b, a), as well as sea ice and glacier schemes which are part of the GEM model itself.

To improve initial conditions of soil moisture, soil temperature and snow depth, the RDRS is coupled with the Canadian Precipitation Analysis (CaPA) (Mahfouf et al., 2007; Lespinas et al., 2015; Fortin et al., 2015, 2018) and the Canadian Land Data Assimilation System (CaLDAS) (Brasnett, 1999; Balsamo et al., 2007; Carrera et al., 2015) as detailed hereafter.

In this section, the main aspects of the global/regional atmospheric reforecasts, as well as the precipitation and surface data assimilation systems are presented. Then, observational datasets used for data assimilation as well as verification metrics used throughout the paper are introduced.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

In order to reduce the risk of discovering major issues with the reanalysis after the fact, GDRS, RDRS, CaPA and CaLDAS configurations are kept as close as possible to documented configurations and currently operational versions of these systems at CCMEP. More details on the configuration of each component are given below.

2.1 Atmospheric model

Both GDRS and RDRS are based on the latest stable version of the Global Environment Multiscale (GEM v4.8-LTS) model (Côté et al., 1998b, a; Girard et al., 2014). Their configurations, as summarized in Table 1 and Table 2, are closely related to the control member of the Global and Regional Ensemble Prediction System (GEPS/REPS) (Li et al., 2008; Lavaysse et al., 2012; Houtekamer et al., 2013; Gagnon and Co-authors, 2015; Lin et al., 2016) as well as the Regional Deterministic Prediction System (RDPS) (Bélair et al., 2005; Mailhot et al., 2006; Caron et al., 2016).

Two configurations of the RDRS were used in this study: RDRS-15 having a horizontal resolution of 15 km (which is based on REPS control member configuration), and RDRS-10 having a horizontal resolution of 10 km (which is based on RDPS configuration). RDRS-15 was nested in a configuration of the GDRS having a resolution of 50 km (GDRS-50), and RDRS-10 in a configuration of the GDRS having a resolution of 39 km (GDRS-39). As will be explained later, the GDRS and RDRS are initialized at both 00 and 12 UTC. The following nomenclature will be used to refer to a specific model configuration and initial time: AAAAZZ-HH, where AAAA refers to the prediction system (either GDRS or RDRS), ZZ to the initial time (either 00 or 12) and HH to the horizontal resolution. For example, RDRS00-10 refers to the 00 UTC run of the 10 km configuration of the RDRS. ZZ and/or -HH will be omitted when obvious from the context or when the statement or conclusion applies independently of either the initial time or the horizontal resolution.

The GDRS-50 features a global uniform latitude-longitude grid with 800×400 meshes of $0.45°$, while the GDRS-39 features a Yin-Yang grid having a uniform resolution of $0.35°$ (having a total of $2 \times 835 \times 303$ meshes). The GDRS features 45 staggered hybrid vertical levels going from the surface up to 0.1 hPa (lid of the model). A 900 s time step is used. RDRS-15 (resp. RDRS-10) features a limited area rotated uniform latitude-longitude grid with 726×556 (resp. 960×1080) meshes. The grid covers North America, adjacent oceans, the whole Arctic ocean and part of Europe and Siberia. A 450 s (resp. 300 s) time step is used. In the vertical, it features 48 (resp. 80) hybrid staggered vertical levels going from the surface up to 10 (resp. 0.1) hPa. In all configurations of the GDRS and RDRS, the first momentum level is at ~ 40 m and the first thermodynamic level is at ~ 20 m.

While similar, configurations of the GDRS and RDRS are adapted to their grid configuration: (a) a few parameters differ as required by the resolution change, (b) additional diffusion is added at the pole of the GDRS-50, (c) in the RDRS-15, an upper-boundary nesting procedure in used (McTaggart-Cowan et al., 2011) in addition to the usual lateral boundary nesting, and (d) in the RDRS-10, a more advanced turbulent kinetic energy (TKE) closure is used to take into account boundary layers clouds (Bélair et al., 2005).

Concerning input geophysical fields (i.e. orography, surface roughness length (except over water), subgrid-scale orographic parameters for gravity wave drag and low-level blocking, vegetation characteristics, soil thermal and hydraulic coefficients, and glaciers fraction), the same input databases and processing methods are used in the GDRS and RDRS as in the operational

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

systems, albeit to produce fields at a different resolution. They are derived from a variety of recent geophysical databases using an in-house software (GenPhysX v2.3.4, Zadra et al., 2008), and are fixed in time.

[Figure 2 about here.]

[Table 1 about here.]

[Table 2 about here.]

Finally, as illustrated in Fig. 2, both the GDRS and the RDRS are launched twice a day (every 12-h at 00 UTC and 12 UTC) and integrated for 24-h in an intermittent cycling manner. Both rely on the 37 pressure level from ERA-Interim as initial atmospheric conditions of the five main variables of GEM. To avoid the effect of initial shock, a diabatic filtering procedure (Fillion et al., 1995) is applied at the beginning of each integration, i.e. during the first 6-h in the GDRS and the first 3-h in the RDRS. Thus, given the resolution jump, the spin-up required for some of the variables such as cloud and precipitation and this filtering procedure, it was decided to avoid using the data for the first six hours following the beginning of the integration for land-data assimilation purposes.

Outputs from GDRS-50 and RDRS-15 are available for a five-year test period (2010-2014), whereas outputs from GDRS-39 are available for the whole period covered by ERA-Interim (1980-2018). Outputs from RDRS-10 will be produced for the same period by CCMEP, and are already available for 2000-2017 (1980-2018 planned).

2.2 Surface assimilation

The land-surface treatment (soil moisture, soil temperature and snowpack conditions) differs notably between the GDRS and RDRS. For the GDRS, as shown in Figs 1 and 2, initial conditions of each simulation are obtained from an a priori offline simulation of the CCMEP land-surface model, GEM-Surf (Carrera et al., 2010; Bernier and Bélair, 2012), directly forced by ERA-Interim atmospheric variables.

For its part, the RDRS is coupled with CaLDAS (Balsamo et al., 2007; Carrera et al., 2015), meaning that the latter is driven by the RDRS which then uses the surface analyses produced by CaLDAS as surface initial conditions for the next integration as illustrated in Fig. 2. CaLDAS uses a one-dimensional Ensemble Kalman Filter (EnKF) to estimate soil moisture and soil temperature, and an optimal interpolation (OI) scheme to estimate snow depth. In addition, considering that precipitation is one of the main driver of surface and soil processes, CaLDAS further relies on the Canadian Precipitation Analysis (CaPA) (Mahfouf et al., 2007; Lespinas et al., 2015; Fortin et al., 2018) system to provide it with a 6-h precipitation analysis as input. CaPA relies on an OI method similar to the one used for snow depth. In all cases, an ensemble of analyses is obtained by perturbing the meteorological inputs to the land-surface scheme, as described by Carrera et al. (2015). The combination of CaPA and CaLDAS configurations designed to initialize the GEM model in its regional configuration is known as the Regional Surface Analysis System (RSAS). All the analyses are performed at the resolution of the background field for a sub-domain covering the whole North America (co-located meshes with RDRS). Furthermore, in order to avoid using data within the spin-up period of the model, only data from the 6-h to 18-h reforecast leadtime serves as background as shown in Fig. 2. Figs 1 and

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

2 sum-up this approach. Each RSAS analysis cycle runs for six hours, because precipitation and snow depth analyses are only available every 6-h (analyses for all others land surface variables are available every 3-h). Hence, following such an approach allows for combining surface observations of temperature, humidity, snow depth and precipitation with the first guess provided by the RDRS and GEM-Surf through various OIs and a one-dimensional EnKF (Brasnett, 1999; Balsamo et al., 2007; Carrera et al., 2015; Fortin et al., 2015).

2.3 Precipitation analysis

[Figure 3 about here.]

As already mentioned, CaPA (Mahfouf et al., 2007; Lespinas et al., 2015) is the system used to produce gridded precipitation analysis for 6-h and 24-h accumulation periods. CaPA combines precipitation observations with a background field obtained from the short-term reforecast provided by the RDRS through an OI method. It aims to correct the error of the latter by spatially interpolating in a transformed space (cubic-root transformation) the differences between the observed values and the corresponding background values at the station locations. Error statistics (standard deviations of background errors and of observation errors, and the characteristic length scale) required to perform the OI are continuously updated based on a temporally adaptive method allowing to implicitly consider seasonal changes in precipitation characteristics (and evolving model error). CaPA also includes an advanced quality control of observations (Lespinas et al., 2015). Although the operational configuration of CaPA assimilates radar quantitative precipitation estimates (Fortin et al., 2015), no radar data is used in this study, due to the availability and the complexity of accessing, processing and controlling the quality of radar data for such a long period in the past.

CaPA has been applied successfully in a number of studies, oftentimes demonstrating great capabilities. See Fortin et al. (2018) for a recent literature review. It is currently operational at CCMEP, providing near real-time precipitation analysis over North America at 10 km resolution and over most of Canada at 2.5 km resolution.

Given that precipitation is one of the most influential forcing of land surface processes, CaPA serves to provide CaLDAS with 6-h precipitation analysis as illustrated in Fig. 1. For that sake, 1-h, 3-h and 6-h precipitation accumulations observations from the SYNOP, SWOB and METAR networks covering the whole North America are assimilated. In order to represent the uncertainty in both the background field and the observations, an ensemble of CaPA 6-h analyses are actually generated in the process (see Carrera et al., 2015, for more details). This precipitation analysis will be later referred to as CaPA-6h.

Several precipitation observation networks, notably the ones that are subject to the most advanced offline quality control and adjustments (see section 2.5) are, however, only available for daily accumulations. As a result, these networks cannot be used directly in the production of the reanalysis as previously described. Hence, in order to take advantage of such networks, CaPA is run again but for 24-h precipitation accumulations in an a posteriori manner based on the first guess provided by either the GDRS or the RDRS. More precisely, in order to prevent from using data in the spin-up period of the model, the 6-h to 18-h forecast leadtime from various sequential integrations are used to build 24-h accumulations valid at 12 UTC as illustrated in

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Fig. 3. The latter then serve as a background field in order for CaPA to produce 24-h precipitation analyses covering the whole domain and period of interest, later referred to as CaPA-24h.

Strict quality control procedures are in place in both CaPA-6h and CaPA-24h to avoid the assimilation of biased observations, and in particular wind-induced undercatch of solid precipitation. If, based on a temperature analysis, solid precipitation is likely to have fallen in the gauge, wind speed observations are used to determine if a precipitation observation should be assimilated. Different wind speed thresholds are used depending on the network, the type of gauge and whether the station is manned or automated (see Lespinas et al., 2015, for more details on the quality control procedure).

An important difference between CaPA-6h and CaPA-24h is that the former is part of CaLDAS and thus contributes to the initialization of the atmospheric model, whereas the latter serves to improve the final precipitation product through a post-processing step that can be launched a posteriori once all integrations of the atmospheric model have been performed. Hence, it is possible to reprocess the final 24-h precipitation product at a small computational cost if more precipitation datasets (or datasets of better quality) become available.

For many environmental modelling applications, a 24-h time step for a precipitation product is too coarse. In order to provide users with realistic hourly precipitation rates and accurate accumulations, the CaPA-24h precipitation analysis has been disaggregated to an hourly time step using a two-step procedure. Firstly, the CaPA-6h is disaggregated to an hourly time step by linearly rescaling the atmospheric model's hourly precipitation rates in order to match the 6-h accumulations estimated by CaPA-6h. In cases where the model precipitation was negligible but precipitation was observed, a constant precipitation rate was assumed for that time period. Secondly, the same procedure is applied to the CaPA-24h precipitation analysis, but using the disaggregated CaPA-6h product as the reference. This same approach has been applied to RDRS-15 and RDRS-10. The resulting datasets are referred to as CaPA-1h on Fig. 1.

2.4 Snow analysis

The snow analysis follows closely the operational method used at CCMEP. All state variables of the ISBA snow model (Bélair et al., 2003a) are cycled, with the exception of snow depth, which is obtained from an external analysis (Brasnett, 1999). This external analysis uses an OI approach (Brasnett, 1999) to blend a first guess of snow depth provided by the ISBA model. The two main differences with the operational CCMEP analysis are that the first guess comes from the ISBA model rather than from a simpler degree-day model and that an ensemble of analyses is produced, one for each CaLDAS member. Random perturbations are added to the precipitation and temperature fields that are provided to the external snow model in order to obtain the background field for the analysis system. No perturbations are added to the snow depth observations themselves.

2.5 Surface observations used for assimilation and verification

In order to produce a reanalysis of precipitation and land surface, in addition to a background field coming from GEM reforecasts, observations are required as input to the assimilation procedure: observations for total precipitation accumulation (1-h, 3-h, 6-h and 24-h), 2 m above ground level absolute and dew point temperatures, 10 m wind speed and snow depth. In the present study and in the forthcoming 1980-2018 reanalysis, it was decided to only rely on surface observations and no remote

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

sensing observations, such as satellite or radar data. Indeed, the latter does not cover the whole period of interest and were put aside mostly in order for the characteristics of the reanalysis to be consistent in time. Furthermore, some datasets are only available for a small number of years. For example, radar data is only used operationally in the precipitation analysis since November 2014. As a result, changes in the observing system will be more incremental and not drastic. Instead, these datasets can be kept for quality control and evaluation purpose.

For sample periods evaluated in the current study, observations from ECCC's operational archive of data used for NWP along with surface precipitation datasets from ECCC's climatic archive are used. Surface observations from ECCC's operational archive include all North American Surface Synoptic Observations (SYNOP), Surface Weather Observations (SWOB), METeorological Aerodrome Reports (METAR), along with other more specific networks: US COOP station data in Standard Hydrological Exchange Format (SHEF) and observations from the Réseau météorologique coopératif du Québec (RMCQ).

These operationally archived datasets cover at most from 1992 to present and thus not the whole period of the final reanalysis product. However, in addition to the pre-operational observational archives from ECCC, it is planned to further rely on the Integrated Surface Data (ISD, DS463.3 - http://rda.ucar.edu/datasets/ds463.3/) (Lott et al., 2001) prior to 2000. This dataset is mostly composed of stations available in ECCC's operational archive and a few additional datasets. However, DS463.3 has been subject to additional off-line automated quality control.

In order to improve the 24 h a posteriori precipitation analysis, supplementary quality controlled observations were pulled from ECCC's climate observations archive. These observations are available as part of the so-called Adjusted Daily Rain and Snow (AdjDlyRS) observations dataset (Wang et al., 2017). These observations were adjusted for systematic errors, and in particular undercatch and evaporation caused by wind effects, gauge-specific wetting loss, as well as for trace precipitation amounts. This recently released dataset which features 3346 stations is considered as one of the most accurate source of retrospective precipitation data in Canada. AdjDlyRS stations are mainly based on manual stations from the Canadian synoptic network, which are known as the most reliable observations. For this reason, stations belonging to the AdjDlyRS dataset are not filtered out by CaPA's quality control for solid precipitation measurements. Observations from all other datasets are subject to this procedure, which has been shown to considerably improve the bias (Lespinas et al., 2015).

[Table 3 about here.]

Table 3 summarizes the surface observation datasets used by CaLDAS and those used by CaPA for the 6-h analysis coupled to CaLDAS (CaPA-6h), as well as for the 24-h a posteriori precipitation analysis (CaPA-24h).

2.6 Additional surface observations used for evaluation purposes

Additional surface observations were used in order to evaluate precipitation, snow, and runoff predicted by the RDRS. For precipitation, two observation-based products were selected. The first product comprises 24-h precipitation accumulations over the United States obtained from the National Centers for Environmental Prediction's Stage IV analysis. The Stage IV analysis is a national product mosaicked from regional multi-sensor (radar+gauges) precipitation analyses (MPEs) produced by the 12 River Forecast Centers of the National Weather Service (Lin and Mitchell, 2005). Its resolution of 4 km is better

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

than that of the RDRS, and thus is expected to provide more details about precipitation over the United States than the RDRS, but it does not cover neither Canada nor Mexico, and is not valid over water. The second product is version 2.3 of the Global Precipitation Climatology Project (GPCP), which is a global monthly product available on a 2.5×2.5 degree grid, and thus covers the whole domain of interest, but at much lower resolution than the RDRS. It merges gauge and satellite data into a seamless product (Adler et al., 2003, 2018). For snow depth, density and water equivalent, a comprehensive Canadian database of snow surveys was selected (Brown et al., 2019).

For the evaluation of runoff, a Canada-USA transboundary watershed was selected: the Lake Erie watershed. Lake Erie, the tenth largest freshwater lake in the world by area, straddles the Canada-US border. It is part of the Laurentian Great Lakes system (Fortin and Gronewold, 2012), receiving water from Lake Michigan-Huron and draining into Lake Ontario through Niagara Falls. Its level has been steadily increasing since 2013, leading to frequent flooding in recent years (Gronewold and Rood, 2019), and is thus the subject of many studies looking at better predicting its water balance, such as the Great Lakes Intercomparison Project for Lake Erie (GRIP-E; Mai et al., 2020b). Its watershed, defined as the land draining between Port Huron, at the outlet of Lake Michigan-Huron, and Niagara Falls, covers and area of approximately 78000 km^2. Note that this watershed includes lands that drain into Lake Erie through the smaller Lake St. Clair. Inflows into Lake Erie from its watershed is not measured directly, but is estimated based on gauged tributaries using the area-ratio method (Fry et al., 2013). Streamflow gauge data were obtained from the United States Geological Survey (USGS) and the Water Survey of Canada (WSC) over the years 2010 through 2014 for 31 gauges located on rivers that drain into Lake Erie. The gauges were selected by Mai et al. (2020b) as part of a hydrological model intercomparison study.

2.7 Verification metrics

Various verification metrics are used throughout the paper to evaluate prediction errors. Let X_n be a prediction, O_n the verifying observation, and $\{\epsilon_n = O_n - X_n, n = 1, \ldots, N\}$ a set of prediction errors. Commonly used metrics include the bias of the error, the standard deviation of the error (or STDE), the root mean square error (or RMSE) and the mean absolute error (or MAE):

$$\begin{aligned}
\text{Bias} &= \frac{1}{N}\sum_{n=1}^{N}\left(\overline{O} - \overline{X}\right) = \frac{1}{N}\sum_{n=1}^{N}\epsilon_n = \overline{\epsilon} \\
\text{RMSE} &= \sqrt{\frac{1}{N}\sum_{n=1}^{N}\epsilon_n^2} \\
\text{STDE} &= \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\epsilon_n - \overline{\epsilon}\right)^2} = \sqrt{\text{RMSE}^2 - \text{Bias}^2} \\
\text{MAE} &= \frac{1}{N}\sum_{n=1}^{N}|\epsilon_n|
\end{aligned} \tag{1}$$

where $\overline{O}$ (resp. $\overline{X}$) is the arithmetic mean of the observations (resp. forecasts). Note that bias is defined as being positive when the mean of the observations is larger than the mean of the predictions. The ideal value for all of these scores is zero.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Streamflow predictions are compared to observations using the Nash-Sutcliffe efficiency (NSE):

$$\mathrm{NSE} = 1 - \frac{\sum_{n=1}^{N}(O_n - X_n)^2}{\sum_{n=1}^{N}(O_n - \overline{O})^2} \tag{2}$$

The ideal value for NSE is one; values above zero indicate better skill than climatology.

Precipitation predictions are compared to observations mainly through categorical scores. For a given threshold T, the number of hits H, misses M, and false alarms F are defined as follow:

$$\begin{aligned} H &= \#\{(O_n > T) \wedge (X_n > T)\} \\ M &= \#\{(O_n > T) \wedge (X_n \leq T)\} \\ F &= \#\{(O_n \leq T) \wedge (X_n > T)\} \end{aligned} \tag{3}$$

where $\#$ denotes the cardinal of the ensemble. From H, M, and F, the probability of detection (POD), the false alarm ratio (FAR), the frequency bias index (FBI), and the equitable threat score (ETS) are defined as follow:

$$\begin{aligned} \mathrm{POD} &= H \cdot (H + M)^{-1} \\ \mathrm{FAR} &= F \cdot (H + F)^{-1} \\ \mathrm{FBI} &= (H + F) \cdot (H + M)^{-1} \\ \mathrm{ETS} &= (H - H_c) \cdot (H + M + F - H_c)^{-1} \end{aligned} \tag{4}$$

where $H_c = (H + F) \cdot (H + M)/N^2$ is the number of hits expected by chance alone. Ideal value for FAR is zero, and one for POD, FBI and ETS.

In addition to these categorical scores, the partial mean of both observations and forecasts are compared in order to assess bias as a function of precipitation intensity. The partial mean (PM) of a sample $\{X_n, n = 1, \ldots, N\}$ for a threshold T is defined as the arithmetic mean of sample values that are smaller or equal to T. As the threshold increases towards the largest value of the sample, the partial mean converges to the arithmetic mean of the sample. The ideal value of PM for a forecast is the partial mean of the observations for the same threshold.

3 Impact of surface assimilation

The impact of surface assimilation on the quality of the reanalysis was assessed using the RDRS-15 configuration (see Table 2). Although this version was run from 2010 to 2014, the evaluations were carried out on much shorter periods, more precisely the 2011 winter and 2011 summer periods, as well as the years 2013-2014. It is worth noting that the 2011 winter and 2011 summer have also been used by CCMEP to evaluate/validate all the forecast and analysis systems which were operational in 2015-2016, thus facilitating the comparison with the systems that were operational at that time.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

In the following subsections, RDRS is evaluated against other reference experiments and datasets. As a general rule, surface analyses are not directly evaluated since all available observations were assimilated. The short-term forecasts that are part of the surface reanalysis creation process are evaluated instead. Evaluation of these short-term forecasts is also important because this data is often used to force offline surface, hydrology and environmental models over long periods. The only exception to this rule is for precipitation. In that case, the precipitation analysis itself is evaluated, using a cross-validation procedure presented in detail by Lespinas et al. (2015).

A main reference in these evaluations is the RDPS, i.e the operational regional weather forecasting system from CCMEP. The horizontal resolution of RDPS has increased over time, but has remained at 10 km since October 2012. Indeed, considering that both systems (RDRS and RDPS) are based on the same atmospheric and surface models (GEM/ISBA), results of the reforecast are expected to be inline with the latter, and in particular should feature similar error signatures. The objective is not to obtain better results with RDRS than with RDPS, since RDPS is initialized with an in-house analysis developed for the latter on its native vertical coordinate and at higher resolution (both vertical and horizontal) and is using more advanced data assimilation method than what was used for ERA-Interim. The objective is rather to obtain forecasts of comparable quality for earlier years, with a system that can be more easily run from 1980 to the present day.

In order to assess the impact of land data assimilation on the quality of the resulting reanalysis, three RDRS-15 configurations are compared to RDPS and to each other: one in which surface initial conditions are provided by CaPA and CaLDAS (RDRS-15, also referred to RDRS+CaLDAS for clarity in the context of this comparison), one in which surface initial conditions are provided by GEM-Surf running in open loop mode and forced by ERA-Interim (RDRS+Open-Loop), and one in which surface initial conditions are cycled, each run starting from a 12 hour forecast of surface conditions obtained from the previous run of RDRS-15 (RDRS+Cycle). Surface initial conditions in RDRS+CaLDAS and RDRS+Open-Loop benefit from the assimilation (either through CaLDAS or ERA-Interim) of surface observations. On the other hand, RDRS+CaLDAS and RDRS+Cycle benefit from the fact that the same atmospheric model is used to drive the land-surface model in analysis and forecast mode. Only RDRS+CaLDAS has both features, but is more costly to operate. RDRS+Open-Loop is closer to the dynamical downscaling approach used in a NWP context (with initial conditions coming from an independent analysis), and RDRS+Cycle is close to the dynamical downscaling approach used for regional climate modelling (with initial conditions obtained from a previous cycle of the limited area model).

Verifications are performed against radiosonde observations as well as surface observations from synoptic stations in North America. Comparing these three systems at the initial integration time would only inform us on the fit to observations of these systems, since these are not independent observations. Forecasts with leadtimes of 6-h to 24-h are compared instead.

3.1 Impact of surface assimilation on the free atmosphere

[Figure 4 about here.]

In this first evaluation, results from the whole atmosphere are evaluated against radiosonde balloon observations that are launched twice a day (00 UTC and 12 UTC) across the whole North American sub-continent in order to assess upper-air

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

conditions, and in particular absolute temperature T, dew point temperature T_d, dew point depression $T - T_d$, wind speed $||U||$, and geopotential height Z_{geop} as a function of atmospheric pressure is also reported. This methodology is standard at CCMEP and represents generally one of the very first steps when evaluating atmospheric systems. The same methodology is thus applied here. It is noteworthy that balloon vertical height measurement has historically been based on the environment pressure. Hence, each balloon's measurement is not located at the same height above ground level (agl) or above sea level (asl), but rather reported at standard pressure levels. As a result, comparisons with model outputs are also done based on pressure levels. While such an approach is convenient and straightforward for the free atmosphere, results obtained close to the surface have to be interpreted with caution: the lowest measurement of each balloon depends on the local surface pressure of its launching location (which is strongly correlated with the height asl). Thus, surface error is not only represented by the lower point of the profiles but spanning across the bottom of the profile (i.e. from 1000 hPa up to 850 hPa).

Fig. 4 a) show comparison results for February/March 2011 and Fig. 4 b) shows results from the same approaches for July/August 2011. Comparisons are done after 24-h integration of the atmospheric model for the whole North America and results from the runs starting at 00 UTC and 12 UTC are combined. In each of the sub-figures the curves on the left show the bias while the curves on the right show the STDE.

GEM integrations that are initialized with surface open-loop results forced by ERA-Interim (blue dashed-dotted line) are compared with integrations initialized with the CaLDAS analysis produced with the coupled system (orange continuous line). In both cases the atmosphere is initialized directly with ERA-Interim. It is noteworthy that the coupled system was run starting in January 2010 in order for the surface fields and notably the root zone water content to be properly initialized. In this figure, the operational RDPS results (2011 final cycles) are also shown as a reference (black dashed line) since the forecast component (GEM) of this system is very similar to the one used to produce the reanalysis.

The CCMEP analysis used to initialized the RDPS is much closer to the upper-air observations than ERA-Interim (not shown here), but this does not translate into a major advantage for the RDPS in the STDE of the forecast after 24 hours of integration, as the weather systems have had time to evolve during the integration.

From Fig. 4 a), it can be seen that during the winter the continuous orange lines and the dotted-dashed blue lines are very close to one another. Hence, results are almost unaffected by surface initialization, and only absolute temperature features small improvements (mostly for the bias) for the coupled reforecast compared to the uncoupled one. In the boundary layer, winter results are also generally in good agreement with the operational RDPS, with the exception of a warm bias and increased STDE of temperature and dew point depression below 850 hPa which are both seemingly caused by initial surface conditions. The whole atmosphere is also wetter (as illustrated by a negative bias in dew point depressions) with a decreased wind modulus while the geopotential height is closer to the observations than the RDPS for both reforecast experiments which is likely to be caused by ERA-Interim initial condition.

During the summer on the contrary, Fig. 4 b), important differences between the two reforecast experiments can be observed in the temperature and dew point depression between RDRS+CaLDAS and RDRS+Open-Loop while both the wind modulus and geopotential height are almost unchanged. The coupled system, which is in very good agreement with the operational RDPS, shows great improvements over the open-loop initialized results, with an impact up to 700 hPa on both the temperature

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

and dew point depression. Indeed, an important error develops in the un-coupled reforecast with a local maximum at 900 hPa. The shape and curvature of the bias for these two variables are also significantly different from the coupled approach (with the latter being more consistent with the RDPS): a warm dry bias develop above 1000 hPa and up to 700 hPa in the uncoupled reforecast. These increased errors could be due to an inconsistency caused by the differences between GEM-Surf driven by ERA-Interim and GEM.

By further exploring these scores in details for the various run hours, leadtimes and geographical regions (not shown here), these errors turned out to be clearly increasing during summer afternoons inland, i.e. in places where convective instabilities driven by the sun heated surface dominate the diurnal cycle. This points out to a deficient liquid water content in the soil at initialization preventing a proper on-set of latent heat flux and evapotranspiration and causing the atmosphere close to the ground to become too dry and hot in the open-loop initialized experiment. Fig. 4 thus clearly illustrates the impact of land-surface and precipitation assimilation on the whole atmosphere.

From this first evaluation, it can be concluded that the coupled reforecast approach leads to atmospheric results which are more inline with the operational model for both seasons. They are always closer to the latter as well as observations across the atmosphere than results from the reforecast initialized by the open-loop. This shows that the proposed approach is not only capable of performing properly in the boundary layer but also in the upper-air, demonstrating that the use of high-quality global atmospheric reanalysis such as ERA-Interim coupled with a land-data assimilation system can alleviate some of the cost of performing a regional reanalysis, since it is not necessary to perform a costly atmospheric assimilation procedure in order to obtain initial atmospheric conditions.

3.2 Impact of surface data assimilation on surface fields

Following the above atmospheric comparisons, this section aims to evaluate the very same winter/summer 2011 experiments, but based on 2 m agl absolute and dew point temperatures as well as wind speed observations from the SYNOP network in North America. Table 4 summarizes the differences between RDRS+Open-Loop 24-h and RDRS+CaLDAS integrations in terms of averaged RMSE changes (in %) for various geographical sub-domains. Cell values are shown in bold when a difference of more than 10% is observed. Negative values, corresponding to cases where RDRS+CaLDAS is worse than RDRS+Open-Loop, are underlined.

At the surface, the coupled reforecast approach almost always shows better results for absolute and dew point temperatures than simulations initialized with the open-loop regardless of the run hour or geographic domain. With the exception of integration starting at 00 UTC for the eastern USA domain in winter, absolute and dew point temperatures from RDRS+CaLDAS are always better by at least a couple of percent, and these improvements go up to 23 % in the eastern part of the USA for the summer. Generally, improvements are larger in summer than in winter. Concerning the wind modulus, differences are most of the time small. Nonetheless, wind modulus for USA domains in the summer show improvements on the order of a couple of percents.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Similar differences are obtained for integrations initialized at either 00 UTC or 12 UTC. However, improvements for dew point temperature tend to be larger in summer for integrations initialized (and valid) at 00 UTC, i.e. in the evening for North America.

[Table 4 about here.]

Instead of relying on initial surface conditions from an open-loop run, a second alternative to RDRS+CaLDAS is considered in which initial conditions are obtained through cycling: in the RDRS+Cycle experiment, GEM forecasts are initialized at the surface from the output of the previous model run, without any form of land-data assimilation. While this setup provides more consistency between the atmospheric and land-surface prediction, there is a risk of surface model drift, especially when running over such a large domain as North America. A longer model spin-up is also required in order for the root-zone water content to be properly initialized. However, if this approach were sufficiently accurate compared to RDRS+CaLDAS, it would be simpler to set up and less costly to run.

This approach was thus applied for a period of more than two years (2013-2015) and results obtained are illustrated based on STDE and bias in Fig. 5. In order to summarize two years of data, monthly averages of bias and STDE are computed, considering forecasts with leadtimes of 6-h to 17-h. These are the leadtimes typically recommended for continuous offline simulation of surface and hydrology models: the first six hours of each forecast is discarded to minimize model spin-up issues, and the next twelve hours from both the 00 UTC and the 12 UTC runs are combined to obtain a continuous hourly forcing dataset.

In that figure, the two plots present the STDE (top curves, solid lines) and the bias (bottom curves, dashed lines) for absolute temperature (top plot) and dew point temperature (bottom plot). The blue curve (resp. orange curve) corresponds to the RDRS+Cycle (resp. RDRS+CaLDAS) experiment. We can see that the coupled approach is always better for both the bias and STDE regardless of the season. The cycling approach also features degraded results in comparison to the RDRS+Open-Loop (not shown here). Most notably, during the summer months, the approach where the surface is cycled shows an important increase of the absolute temperature bias along with an important decrease of dew point temperature bias. This tends to point out that there is not enough water available in the soil for evaporation causing a warmer/dryer atmosphere close to the ground. Indeed, at these resolutions and when relying on the cold-start/diabatic filter for initialization, GEM requires a spin-up period of at least 6-h for its precipitations to reach a stable level. This deficit of water can be partly explained by the fact that the surface scheme implemented in GEM-Surf, i.e. ISBA, is recognized not to properly retain soil moisture (Alavi et al., 2016). This model deficiency is compensated by CaLDAS through positive increments in soil moisture. During winter, while results from the coupled approach tend to be closer to the observations for both bias and standard deviation, differences are not as important as for the summer season.

Results presented in the above two sections clearly show the added value of the coupled approach with regards to the surface layer and atmosphere. Results obtained are inline with the operational model in the atmosphere. In the surface layer, significant improvements are obtained in comparison with the simpler and more standard open-loop or cycling approaches. It is based on these results that the production of a 10 km configuration of the RDRS, coupled with CaLDAS, was launched at CCMEP for

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

the period 1980-2018. The next section presents an evaluation of this product for surface variables, precipitation accumulations and streamflow.

[Figure 5 about here.]

4 Evaluation of the continuous hourly forcing dataset over 2011-2017

The impact of surface and precipitation analysis on reforecasts has been evaluated on short time periods in the previous section. An improvement of the reforecast is observed when the RDRS is coupled with CaPA/CaLDAS analysis systems. The present section aims to evaluate various aspect of the reanalysis over a longer time period (2011-2014 for RDRS-15 and 2011-2017 for RDRS-10) in order to better appreciate its quality. Comparison against screen-level absolute and dew point temperatures, wind speed, precipitation and streamflow in-situ observations are presented, followed by a comparison of precipitation totals with other products available for North America.

Furthermore, the evaluation is focused on the continuous hourly forcing dataset recommended for offline simulation of surface and hydrology models, and which will be distributed to stakeholders as the main outcome of this project. For absolute temperature, humidity, radiation, wind speed and surface pressure, this dataset is obtained by combining RDRS forecasts with leadtimes of 6-h to 17-h from both the 00 UTC and the 12 UTC runs. For precipitation, the CaPA-24h product, disaggregated to a hourly time-step is used (see section 2.3). The resulting datasets are referred to in this section simply as RDRS-15 and RDRS-10, or else as RDRS-15 + CaPA-24h and RDRS-10 + CaPA-24h when precipitation or offline hydrological simulations are evaluated.

4.1 Absolute temperature, dew point temperature and wind speed

This section focuses on the evaluation of the RDRS-15 and RDRS-10 short-term forecasts (with leadtimes of 6-h to 17-h) against absolute and dew point temperatures observations from synoptic stations in North America. Initial surface conditions for the RDRS-15 and RDRS-10 are provided in all cases by CaLDAS. The performance of the RDPS which was operational at the time is also shown as a reference.

Fig. 6 presents the time evolution of the RMSE for the RDRS-15 (continuous orange line), RDRS-10 (dashed-dotted blue line) and RDPS products (black dashed line), assessed against all SYNOP observations located in North America. A monthly time interval is used to compute error statistics. A clear annual cycle is observed, with errors being generally larger in winter and spring, and smaller in summer and fall. Errors in absolute temperature forecasts (panel a) and dew point temperature forecasts (panel b) show similar patterns. While RDRS-15 and RDRS-10 errors behave similarly, RDPS errors are larger for the earlier years. Focusing on dew point temperature forecast errors (panel b), it can be seen that the differences between RDPS and RDRS errors become smaller starting in 2015 and are very similar after mid-2016. For wind speed (panel c), RDPS and RDRS-15 errors are very similar until the end of 2014, and slightly smaller in winter for RDPS after this date.

The reduction of errors in RDPS forecasts (when compared to both RDRS-15 and RDRS-10) at the end of 2014 and again in mid-2016 is likely the consequence of changes to the operational version of RDPS. In fact, a major upgrade to the data

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

assimilation system was implemented on November 18, 2014 (Caron et al., 2015) and a major upgrade to the GEM model was implemented on September 7, 2016 (Caron et al., 2016). After this date and until the end of the evaluation period, the RDPS relied on the same version of GEM as the RDRS.

When comparing RDRS-15 and RDRS-10 over the whole North America, it can be noted that the 15 km configuration of the system has slightly smaller errors than the 10 km configuration for absolute and dew point temperatures, and slightly larger errors for wind speed, although the differences are small in all cases, and typically smaller than the differences between RDRS-15 and RDPS errors.

[Figure 6 about here.]

A regional evaluation of the performance of the RDRS is presented in Table 5, which summarizes the average changes in RMSE between RDRS-10 and RDRS-15 as well as between RDRS-10 and the operational RDPS by seasons and geographical domains. Scores are again computed against observations from the SYNOP network and based on forecasts with 6-h to 17-h leadtime. The time period for the evaluation is restricted to 2011-2014, during which outputs from all three systems are available.

It can be seen that over this period, i.e. prior to the upgrade of the atmospheric data assimilation system and GEM model version of the RDPS, RDRS-10 improves upon RDPS in most regions and for most variables, notably for Alaska and the Canadian Arctic, as well as Western USA. Absolute and dew point temperatures forecasts are improved for all seasons and all domains, with the sole exception of absolute temperature in spring and summer for the Eastern Canada domain. This degradation is caused mainly by errors in the sea surface temperature (SST) obtained from ERA-Interim: indeed, forecasts of similar skill are obtained for RDPS and RDRS when using operational CCMEP SST analyses (not shown). Improvements of more than 10% are obtained in many cases. They are shown in bold in Table 5. For example, a 24% reduction of forecast error is obtained for dew point forecasts over Mexico during the fall season, and a 25% reduction of forecast error is obtained for absolute temperature forecasts over Western USA during the winter season. RDRS-10 wind forecast errors are slightly better than RDPS forecasts for all domains in the summer and fall seasons, with the exception of Mexico during the summer season. During winter and spring, RDRS-10 wind forecasts are slightly better for the USA, and in particular for the western half of the USA, but slightly worse for Canada and Mexico. In all cases, differences in wind speed forecast errors do not exceed 2%. When comparing RDRS-10 and RDRS-15 forecasts, it can be seen that RDRS-10 absolute and dew point temperatures forecasts are generally worse, with the exception of dew point temperature forecasts over Mexico. On the other hand, RDRS-10 wind forecasts are better than RDRS-15 forecasts for all domains and seasons. All of the differences between RDRS-10 and RDRS-15 RMSE are fairly small. The most important RMSE change being -4% for dew point temperature (USA West domain, Winter season), -6% for absolute temperature (Canada East domain, Spring season), and $+4\%$ for wind speed (Canada East domain, Winter season).

[Table 5 about here.]

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Overall, this evaluation suggests that RDRS-10 near-surface forcing data should be competitive with recent RDPS forcing data for driving surface, hydrology and environmental models, at least for absolute temperature, dew point temperature and wind speed.

Although prediction skill is important for a reanalysis dataset, it is also important to assess the diurnal reanalysis bias, which is interesting by itself and also because it affects the RMSE of the reanalysis. The diurnal STDE is also interesting to assess, as it is equivalent to the RMSE of forecasts from which the bias has been removed. Fig. 7 shows, for each season, the diurnal cycle of the bias and standard deviation of the RDRS-10 and RDPS errors assessed against all SYNOP observations located in North America. Each graph is based on seven years of data (2011-2017). The first row of plots presents results for absolute temperature, and the second row presents results for dew point temperature. Each column corresponds to a different season. Each plot contains four curves: two for RDPS forecasts (in green) and two for RDRS-10 forecasts (in orange). The top two curves (filled lines) on each graph correspond to the STDE as a function of time of day (in UTC hours), whereas the bottom two curves (dashed lines) correspond to the bias. Model runs initialized at 00 UTC (resp. 12 UTC) are used for hours 6 to 15 (resp. 0 to 3 and 18 to 21), with results shown using filled (resp. empty) markers.

[Figure 7 about here.]

In terms of standard deviation, RDRS-10 forecast errors are generally smaller than RDPS forecast errors, especially for 00 UTC forecasts valid from 06 to 15 UTC (filled markers), hence during the night and early morning in North America. The error signature of both systems is fairly similar. In particular, errors are larger in winter than in summer, they are larger during the night and early morning during winter (forecasts issued at 00 UTC, filled markers), and generally larger during the afternoon and evening in summer (forecasts issued at 12 UTC, empty markers). The bias signature of both systems is similar for absolute temperature (first row of plots). The RDRS-10 winter absolute temperature bias is better than that of the RDPS for all hours of the day. For other seasons, the bias is slightly worse during the afternoon and evening hours and comparable during the night and early morning. For dew point temperature, the diurnal cycle of RDRS-10 and RDPS is very similar, but the RDRS-10 bias value is always higher than that of the RDPS, which is an indication that the land-surface is wetter in the RDRS-10. During the night and early morning, this always corresponds to an improvement in the bias. During the afternoon and evening hours, the RDRS-10 bias is generally worse than the RDPS bias. Hence, environmental models that are particularly sensitive to humidity might behave differently when forced with RDRS-10 data compared to RDPS data.

4.2 Precipitation

Two strategies were considered in order to assess the precipitation reanalysis skill and bias at both 15 and 10 km: (1) seasonal accumulation maps were obtained for the CaPa-24h reanalysis at both resolutions and visually compared to Stage IV MPE (Lin and Mitchell, 2005) and GPCP (Adler et al., 2003, 2018) for each season of the period 2010-2014, and (2) verification scores routinely used for evaluating the operational CaPA precipitation analysis were computed over the same period for both reanalysis products, but also for the background field used for each product.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

4.2.1 Seasonal accumulation

Figure 8 compares the seasonal accumulations of precipitation for the two versions of the CaPA-24h reanalysis (at 15 and 10 km resolution) to the seasonal accumulations of precipitation estimated from Stage IV MPE and GPCP datasets. Each column corresponds to a season: Spring (MMA), Summer (JJA), Fall (SON) and Winter (DJF). Each row presents accumulations for a different product: a) CaPA-24h based on RDRS-10, b) CaPA-24h based on RDRS-15, c) Stage IV MPE, and d) GPCP. Of the four products, only GPCP provides global coverage. However, it can be seen that the domain covered by both versions of CaPA covers all of North America, including the coastal oceans. Stage IV MPE is limited to the conterminous US, plus the Columbia river basin in British Columbia and the Rio Grande basin in Mexico.

Over the oceans, large differences are observed between the three products providing coverage: CaPA-24h is nearly identical to its background field, since very few observations are available to constrain the analysis. Thus, it is essentially the RDRS precipitation accumulations that are shown over the oceans in the first two rows of Fig. 8. It appears that the RDRS-15 precipitation matches better the GPCP precipitation, but both RDRS-15 and RDRS-10 seem to overestimate precipitation, in particular over the Gulf Stream. Difference over the ocean between RDRS-10 and RDRS-15 can be explained by the use of differing shallow convection parametrization: while the latter is closer to observation, it breaks conservation laws (Bélair et al., 2005). In the meantime, the overestimation of both approaches might be indicative of an excess of evaporation at mid-latitude in the GEM model, leading in turn to an excess of precipitation, as suggested by the evaluation performed by Deacu et al. (2012) over Lake Superior.

Over the USA, Stage IV and GPCP agree well for the larger scale features, with Stage IV MPE providing more details. Even at 15-km, CaPA-24 also provides much more details than GPCP, and generally agrees well with Stage IV MPE. However, CaPA-24h clearly lacks precipitation during the summer months in the Southeast of the US. This is an area known for its strong convective storms in summer, and the associated precipitation is likely not well captured in neither the GEM model background field nor in the in-situ observational dataset assimilated by CaPA. Over Canada, significant differences are observed between GPCP and CaPA-24, both in terms of precipitation accumulations and variability. This is expected, as neither product relies on a dense precipitation network (especially in northern Canada). Furthermore, the CaPA reanalyses assimilate datasets that are not considered by GPCP (such as the RMCQ network in the Province of Québec), as well as precipitation observations that have been corrected for known biases (the AdjDlyRS dataset).

[Figure 8 about here.]

4.2.2 Scores

In addition to a visual comparison of CaPA-24h reanalyses with reference precipitation products available over North America, an objective verification was also performed. Evaluating precipitation analyses is however more challenging than evaluating precipitation forecasts, since independent verification data is hard to come by, especially in regions with already sparse networks like northern Canada. One option is to rely on data-denial experiments. For CaPA operational analyses and reanalyses, leave-one-out cross-validation is routinely used, as it is a by-product of the in-line quality control procedure (Lespinas et al.,

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

2015). Verification metrics generated by this procedure include the Probability of Detection (POD), the False Alarm Ratio (FAR), the Frequency Bias Index (FBI), the Equitable Threat Score (ETS). It is also useful to compare the Partial Mean (PM) of the observations with that of the analysis in order to assess quantitative bias as a function of precipitation thresholds.

These different verification metrics are presented for each season in Fig. 9. Each row corresponds to a season (DJF, MAM, JJA and SON) and each column to a verification metric. The x-axis is the precipitation threshold (in mm day^{-1}) for which the verification metric is computed, and the y-axis corresponds to the value of that metric. For POD, FAR, FBI and ETS, four curves are shown on each graph. Continuous (resp. dashed) lines present the verification metrics for the reanalyses (resp. background field). Blue (resp. orange) curves correspond to the 15 (resp. 10) km products. In the last column, the partial mean of the observations is shown in black. It corresponds to the ideal value for the PM verification metric. For FBI, the ideal value of one is also highlighted using a black line.

As these are generally considered more accurate, verification is performed against manual synoptic stations in the US and Canada, plus climate stations from the AdjDlyRS dataset. Observations that do not pass the quality control of CaPA are discarded from the verification dataset in all cases.

It is clear from this figure that both analyses are more skillful (in terms of ETS) and less biased (both in terms of FBI and PM) than their background field. False alarms (FAR) are also significantly reduced in the reanalysis for all seasons and thresholds. The probability of detection (POD) in winter for thresholds of 2 $\mathrm{mm/24h}$ or less is the only score for which the reanalysis is worse than the background field. However, given the improvements in the FAR, FBI and ETS, this is deemed acceptable.

What is less clear is the added value of the 10 km compared to the 15 km products: the POD is similar or slightly improved, but the FAR is degraded, in particular for larger precipitation thresholds. The FBI is higher at 10-km resolution, which results in a degraded categorical bias for lower thresholds, but an improved categorical bias for higher thresholds. The ETS is almost always slightly worse at 10-km resolution. The PM is nearly identical at both resolutions. In summary, the 10-km is slightly less skillful than the 15-km reanalysis (as measured by the ETS), and this is due to an increased ratio of false alarms.

In general, the degradation in the FAR and ETS of the analysis when going from 15- to 10-km is not seen in the background field (with the notable exception of higher precipitation thresholds in summer). Furthermore, in most cases the difference in the scores of the two reanalyses is larger than the difference in the scores of the two background fields, which tends to be small. In spring, summer and fall, the POD, FBI and PM of the background field are however generally higher at 10 km, reflecting a worsening of the positive precipitation bias already present at 15 km.

[Figure 9 about here.]

4.3 Snowpack

Snow depth and snow water equivalent (SWE) are of interest for surface and hydrological prediction. They are both routinely observed, but fewer automatic SWE measurements are available. Furthermore, SWE is often reported only around the first and fifteenth day of each month. In order to be able to compare the skill and bias of these two properties, verification was restricted to cases where both quantities were available in the Canadian historical snow survey database compiled by Brown et al. (2019). Only snow surveys taken at an altitude within 300 m of the atmospheric model topography were considered. The evaluation

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

was performed over January 1 – May 15, 2014, November 15 – January 31, 2015, since no observations were available during the warm season. Pairs of observations and corresponding analysis values were obtained at each site for the first and fifteenth day of each month, with a tolerance of ± 6 days on the observation date in order to maximize the number of measurements used for the evaluation while avoiding to over-represent in the verification dataset stations reporting more often than once every two weeks. Indeed, manual snow survey sites are typically visited at most twice per month. Bias, STDE, MAE and RMSE were then computed for RDRS-10. While satisfactory results were obtained for SWE and snow depth, the performance was rather poor for snow density.

A code review revealed that an error had been introduced in the GEM model when implementing the parameterization for maximum snow density proposed by Bélair et al. (2003a). Indeed, equations 21 and 22 in Bélair et al. (2003a) are valid for snow depth expressed in cm, whereas snow depth is expressed in m in both the paper and the GEM code. In addition to being present in the reanalysis, this bug impacts all operational NWP systems based on GEM as of 2020. It will be corrected in future releases. In order to assess the impact of the bugfix, the surface reanalysis was re-launched for 2014.

Table 6 presents the bias, STDE, MAE and RMSE of snow density, snow depth and SWE with and without this bugfix. When using the correct implementation of the maximum snow density parameterization, the bias on snow density drops from 144 to only 7 $\mathrm{kg\ m^{-3}}$ and STDE is also lowered from 90 to 79 $\mathrm{kg\ m^{-3}}$. Consequently, RMSE is reduced from 169 to 78 $\mathrm{kg\ m^{-3}}$. MAE is also considerably reduced from 155 to 57 $\mathrm{kg\ m^{-3}}$. Modest improvements in snow depth predictions are however observed for all metrics: differences range from a 2 cm improvement for bias, STDE and MAE to a 3 cm improvement for RMSE. With the bugfix, the bias of snow depth predictions is -10 cm, STDE is 16 cm, MAE 12 cm and RMSE 19 cm. The impact on the bias and skill of SWE predictions is more important than for snow depth, but not as important as for density. For SWE, the positive bias of 11 $\mathrm{kg\ m^{-2}}$ obtained with the bug becomes negative and slightly worse (-20 $\mathrm{kg\ m^{-2}}$) with the bugfix. STDE is however improved from 67 to 43 $\mathrm{kg\ m^{-2}}$. The net effect on RMSE is a reduction of the error from 68 to 54 $\mathrm{kg\ m^{-2}}$. MAE is also improved with the bugfix from 52 to 34 $\mathrm{kg\ m^{-2}}$.

[Table 6 about here.]

4.4 Streamflow

In a first attempt to assess the suitability of the dataset for hydrological prediction, the GEM-Hydro model (Gaborit et al., 2017) was used to predict the inflows into Lake Erie. The land-surface of the GEM-Hydro model was set up on a 10 km grid, the routing was performed on a 1 km grid, and both were calibrated using the RDRS-15 + CaPA-24h forcing, using a methodology described in Mai et al. (2020b), over the years 2011-2014.

[Figure 10 about here.]

Figure 10 shows a comparison of the simulated inflows obtained when forcing the calibrated GEM-Hydro version with three different datasets: RDRS-15 + CaPA-24h, RDRS-10 + CaPA-24h and RDPS + CaPA-24h. Panel (a) shows the daily inflows from 2011 to 2015 obtained by aggregating the flows from all of the tributaries draining directly into Lake Erie. Panel

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

(b) presents the same data as in panel (a) but ranked for lowest to highest value, and using a log-scale. Panel (c) shows the cumulative runoff over this five year period obtained for the observations and the model by summing up the daily inflows, dividing by the area of the watershed, and converting units into meters.

The comparison between observed and modelled daily inflows (panel a) shows that GEM-Hydro is able to represent the dynamics of the watershed in periods of both low and high inflows for all three sets of forcing data. Nash-Sutcliffe Efficiency (NSE) over 2011-2014 (2010 was used for model spinup) for the three experiments are respectively 0.77 for RDRS-10 + CaPA-24h, 0.78 for RDRS-15 + CaPA-24h and 0.77 for RDPS + CaPA-24h. It is not surprising to obtain a slightly better NSE with the RDRS-15 forcing, since they were used to calibrate the model. Panel (b) shows that both low and high flows are negatively biased in all three simulations, even if the RDRS-10 + CaPA-24h dataset leads to slightly higher flows at both ends of the spectrum. The comparison between the observed and modelled cumulative runoff (panel c) is useful in assessing the bias of the simulations. It can be seen that better results are obtained in terms of cumulative runoff with the RDRS-10 + CaPA-24h forcing, followed by RDRS-15 + CaPA-24h. RDPS + CaPA-24h leads to more biased predictions, on the order of 17%, compared to 12% for RDRS-15 + CaPA-24h and 8% for RDRS-10 + CaPA-24h. This analysis of simulated streamflow hence provides some evidence that the forcing data based on the RDRS can be useful for hydrological prediction, and in particular for transboundary watersheds of North America. Remaining simulation errors can be partly explained by the presence of large agricultural and urbanized areas in the watershed (Mai et al., 2020b). In particular, tile drains and impervious areas are not represented accurately in the model.

5 Discussion and conclusion

This paper presents and evaluates a methodology for minimizing the computational cost and the data processing burden associated with the production of an atmospheric reanalysis, and still obtain consistent, high-quality and high-resolution atmospheric, surface and precipitation fields. A final precipitation analysis is then obtained in a post-processing step in order to assimilate additional precipitation observations which significantly improves the quality - and local relevance - of the precipitation analysis. The methodology is used to obtain a first reanalysis for the GEM model, which is at the heart of the operational NWP system used by the Meteorological Service of Canada. The reanalysis provides seamless coverage for all of North America and the Arctic Ocean at 10 km resolution and hourly time step, and is of particular interest for hydrological applications in transboundary watersheds sitting on the Canada-US border. Results obtained when forcing a hydrological model in order to predict inflows into the Lake Erie watershed are encouraging.

5.1 Impacts of the land-data assimilation system on the reanalysis quality and production strategy

The various numerical experiments have shown the important role played by the land-data assimilation system in the quality of the product not only at the surface, but also for the boundary layer. Indeed, using CaLDAS allows for better results for absolute and dew point temperatures in the boundary layer (cf. Fig. 4) and at the surface (cf. Fig. 5). Initialized by CaLDAS, the RDRS configuration of the GEM model provides better performance than the operational RDPS configuration (cf. Figs. 6 and 7).

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Unfortunately, the use of CaLDAS increases the computational cost, and forces the production to proceed sequentially, thus limiting the ability to speed up production through parallel processing to the point where sequential production from 1980 to the present time would simply not have been feasible. It was thus decided as a compromise to launch the production in four-year batches starting at the beginning of the hydrological year (September), with the first year being only used for spin-up and thus discarded.

5.2 Skill, bias, local relevance and consistency of the precipitation reanalysis

One of the advantages of the methodology proposed in this paper is that the precipitation observations can both influence the surface and atmospheric fields (through the online Kalman filter land-data assimilation system), and contribute to the final, offline, optimal interpolation of precipitation data. The offline precipitation analysis can include more data (and in particular 24-h accumulations) and its computational cost is small in comparison with the rest of the reanalysis process. The final analysis can therefore be updated in post-production as more data (or data of better quality) becomes available, and thus provide a better fit to specific in-situ observations. Being faithful to specific in-situ observations can be of critical importance for many hydrological modelling applications, for example when simulating past flood events.

The offline OI technique does have limitations, despite having less bias and more skill than the short-term precipitation forecasts provided by the RDRS and used as the background field for the analysis. Indeed, since only the precipitation field is modified, this creates inconsistencies with other variables, such as radiation and humidity, which can be detrimental for some land-surface modelling applications.

Another issue with the OI technique is the assumption that the background field is unbiased, a hypothesis not verified for RDRS precipitation which has been shown to be positively biased. As distance to in-situ stations increases, the OI-based precipitation analysis converges to this biased background field, and is thus locally biased. A bias correction could be applied to the background field before performing the OI, but this requires to first obtain the GEM model climatology. This option will thus be evaluated once the production of the first version of the reanalysis is completed for 2000-2018.

Finally, it was disappointing to realize that the precipitation reanalysis was not of better quality at 10 km than at 15 km. The 10-km background field showed a similar skill but more bias, and the 10-km reanalysis showed slightly less skill and a similar bias. It should not be surprising that precipitation bias might be degraded when model resolution is increased, as numerous factors can adversely impact bias of precipitation forecasts in a NWP system, and in particular the choice and tuning of sub-grid convection and condensation parameterizations. The lack of improvement in model skill with resolution could be real since the difference in horizontal resolution between both RDRS configurations is not huge, but it could also be that the synoptic and climatic networks used for verification are not dense enough to capture improvements to precipitation fields. Finally, the fact that a more skillful analysis is obtained when using a lower resolution background field is not uncommon when using an OI method: a smoother background field can lead to innovations that are more correlated in space and thus more efficiently interpolated.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

5.3 Hydrological applications of the reanalysis dataset

The 15-km version of the reanalysis (RDRS-15) was initially released in 2017 by ECCC and has since been used in multiple hydrological studies. For example, in a paper published by Abaza et al. (2020), the precipitation and temperature forcing data were used to initialize an ensemble hydrological forecasting system for Lake Champlain and Richelieu River. In open-loop mode, the hydrological model represented very well the observed streamflow at the outlet of the lake, with a NSE of 0.92 obtained for the model verification period. The same dataset was also used for a hydrological model intercomparison study performed over the Lake Erie watershed (Mai et al., 2021). Seventeen models of different complexity (including traditional hydrological models as well as machine learning models) were evaluated against 53 streamflow gauges (46 being used for calibration and 7 for validation). A more in-depth evaluation of the performance of the HYPE hydrological model was furthermore presented by Awoye et al. (2019). All models produced satisfactory simulations. In particular, the machine learning models provided excellent performance, and models calibrated at individual stations performed surprisingly well at validation sites. A limitation of this study is that only five years of data were available from RDRS-15. A follow-up study, relying on the 2000-2017 RDRS-10 dataset, is already underway.

5.4 Known issue with snow depth

As explained previously, a bug was identified during the production of the 2000-2017 RDRS-10 reanalysis dataset. This bug impacts all operational configurations of the GEM model, as well as the RDRS-15 reanalysis dataset. It has since been corrected, and the production of the reanalysis for the years 1980-1999 will include the bugfix. This will cause a sudden shift in the bias and skill of the snow depth, density and water equivalent in January 2000. Based on the assessment of the impact of this bug over a one year period, modest changes are expected for snow depth, but snow density and water equivalent will be more impacted. For this reason, the snow density and water equivalent fields will not be distributed for years 2000-2017.

5.5 Dependency on ERA-Interim atmospheric reanalyses

The methodology used for the reanalysis dataset proposed in this paper requires the use of upper-air fields provided by a lower resolution global reanalysis. Although minimal information is required - the data is only used at initial time (00 and 12 UTC) and only for upper-air atmospheric variables - it creates a dependency on an external dataset. When the production of the 2000-2017 dataset started, it was decided to rely on ERA-Interim, which has since been discontinued. For the years 1980-1999 as well as 2018, production will likely continue with ERA-Interim. Tests are currently being performed on 2018 to assess the impact of switching to a different data source for 2019-present. Candidate upper-air fields include ERA5 and CCMEP operational upper-air analyses.

5.6 Future developments

Given the strong demand for the RDRS reanalysis data in Canada ahead of the release of the dataset, we expect that the product will be regularly updated in order to provide data for recent years as well as to take advantages of future improvements in

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

each component of the RDRS, namely CaLDAS, CaPA and GEM. We also expect that the strengths and weaknesses of the dataset that will be identified through ongoing and future hydrological modelling applications of the dataset will guide future developments of the reanalysis. Finally, it should be noted that although the reanalysis dataset presented in this paper was designed with hydrological applications in mind, other uses are now envisioned, including the reconstitution of high-impact events through further dynamical downscaling of the reanalysis, the establishment of a GEM model climatology to derive anomaly forecasts from RDPS outputs, and the use of the reanalysis for driving atmospheric chemistry and air quality models in support of health studies and air emission regulation in Canada.

Data availability. selected output variables from RDRS-15 and RDRS-10 are available through the Canadian Surface Prediction Archive (CaSPAR, https://caspar-data.ca, Mai et al. (2020a)). Additional variables, including upper-air fields, are available upon demand from CCMEP.

Author contributions. Main contributions from each co-authors are as follows: under the supervision of V. Fortin, N. Gasset proposed the initial design of the RDRS-15 as well as a more production friendly and updated design for RDRS-10 and was responsible for the production and evaluation of the RDRS-15 dataset. M. Dimitrijevic oversaw the production and evaluation of the CaPA-24h precipitation reanalysis and performed snowpack evaluation. M. Carrera and B. Bilodeau contributed to the CaLDAS component. R. Muncaster provided the offline land-surface prediction system used for initializing the GDRS. E. Gaborit performed the hydrological evaluation of RDRS over Lake Erie. G. Roy contributed the CaPA component. N. Pentcheva was responsible for the production and dissemination through CaSPAR of the RDRS-10 dataset. M. Bulat was responsible for the preparation of the in-situ datasets used by CaPA and for the production of the CaPA-24h dataset. X. Wang was responsible for the preparation of the in-situ datasets used by CaLDAS and for the evaluation of the near-surface variables. R. Pavlovic was responsible for overseeing the RDRS-10 production, and in particular, for optimizing the allocation of computer resources. F. Lespinas and D. Khedaouiria contributed to the configuration and evaluation of the CaPA-24 precipitation analysis. Finally, V. Fortin and N. Gasset contributed equally to the writing of the paper, with contributions from co-authors, and in particular from M. Dimitrijevic.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The financial contribution of the International Joint Commission to this research project, through its International Watersheds Initiative, is gratefully acknowledged. We wish also to thank Normand Gagnon who provided valuable input for the initial design of the RDRS.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

References

Abaza, M., Fortin, V., Étienne Gaborit, Bélair, S., and Garnaud, C.: Assessing 32-Day Hydrological Ensemble Forecasts in the Lake Champlain–Richelieu River Watershed, Journal of Hydrologic Engineering, 25, 04020 045, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001983, 2020.

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., and Nelkin, E.: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present), Journal of Hydrometeorology, 4, 1147–1167, https://doi.org/10/cb5d4s, 2003.

Adler, R. F., Sapiano, M. R. P., Huffman, G. J., Wang, J.-J., Gu, G., Bolvin, D., Chiu, L., Schneider, U., Becker, A., Nelkin, E., Xie, P., Ferraro, R., and Shin, D.-B.: The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation, Atmosphere, 9, https://doi.org/10.3390/atmos9040138, https://www.mdpi.com/2073-4433/9/4/138, 2018.

Alavi, N., Bélair, S., Fortin, V., Zhang, S., Husain, S. Z., Carrera, M. L., and Abrahamowicz, M.: Warm Season Evaluation of Soil Moisture Prediction in the Soil, Vegetation, and Snow (SVS) Scheme, Journal of Hydrometeorology, 17, 2315–2332, https://doi.org/10.1175/JHM-D-15-0189.1, 2016.

Albergel, C., Dorigo, W., Reichle, R. H., Balsamo, G., de Rosnay, P., Muñoz-Sabater, J., Isaksen, L., de Jeu, R., and Wagner, W.: Skill and Global Trend Analysis of Soil Moisture from Reanalyses and Microwave Remote Sensing, Journal of Hydrometeorology, 14, 1259–1277, https://doi.org/10.1175/JHM-D-12-0161.1, 2013.

Awoye, O. H. R., Bajracharya, A. R., Stadnyk, T., and Asadzadeh, M.: Is the physical hydrologic model HYPE well suited for the simulation of water quantity in North-American watersheds? - A modelling experiment with the newly developed RDRS meteorological reanalysis data, in: AGU Fall Meeting Abstracts, vol. 2019, pp. H33M–2162, 2019.

Balsamo, G., Mahfouf, J.-F., Bélair, S., and Deblonde, G.: A Land Data Assimilation System for Soil Moisture and Temperature: An Information Content Study, Journal of Hydrometeorology, 8, 1225–1242, https://doi.org/10.1175/2007JHM819.1, 2007.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: A Global Land Surface Reanalysis Data Set, Hydrology and Earth System Sciences, 19, 389–407, https://doi.org/10.5194/hess-19-389-2015, 2015.

Bélair, S., Brown, R., Mailhot, J., Bilodeau, B., and Crevier, L.-P.: Operational Implementation of the ISBA Land Surface Scheme in the Canadian Regional Weather Forecast Model. Part II: Cold Season Results, Journal of Hydrometeorology, 4, 371–386, https://doi.org/10.1175/1525-7541(2003)4<371:OIOTIL>2.0.CO;2, 2003a.

Bélair, S., Crevier, L.-P., Mailhot, J., Bilodeau, B., and Delage, Y.: Operational Implementation of the ISBA Land Surface Scheme in the Canadian Regional Weather Forecast Model. Part I: Warm Season Results, Journal of Hydrometeorology, 4, 352–370, https://doi.org/10.1175/1525-7541(2003)4<352:OIOTIL>2.0.CO;2, 2003b.

Benedict, I., Van Heerwaarden, C., Weerts, A., and Hazeleger, W.: The benefits of spatial resolution increase in global simulations of the hydrological cycle evaluated for the Rhine and Mississippi basins, Hydrology and Earth System Sciences, 23, 1779–1800, https://doi.org/10.5194/hess-23-1779-2019, 2019.

Bernier, N. B. and Bélair, S.: High Horizontal and Vertical Resolution Limited-Area Model: Near-Surface and Wind Energy Forecast Applications, Journal of Applied Meteorology and Climatology, 51, 1061–1078, https://doi.org/10.1175/JAMC-D-11-0197.1, 2012.

Bosilovich, M. G., Chen, J., Robertson, F. R., and Adler, R. F.: Evaluation of Global Precipitation in Reanalyses, Journal of Applied Meteorology and Climatology, 47, 2279–2299, https://doi.org/10.1175/2008JAMC1921.1, 2008.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Bosilovich, M. G., Robertson, F. R., and Chen, J.: Global Energy and Water Budgets in MERRA, Journal of Climate, 24, 5721–5739, https://doi.org/10.1175/2011JCLI4175.1, 2011.

Brasnett, B.: A Global Analysis of Snow Depth for Numerical Weather Prediction, Journal of Applied Meteorology, 38, 726–740, 1999.

Brown, R., Fang, B., and Mudryk, L.: Update of Canadian Historical Snow Survey Data and Analysis of Snow Water Equivalent Trends, 1967–2016, Atmosphere - Ocean, 57, 149–156, https://doi.org/10.1080/07055900.2019.1598843, 2019.

Bélair, S., Mailhot, J., Girard, C., and Vaillancourt, P.: Boundary Layer and Shallow Cumulus Clouds in a Medium-Range Forecast of a Large-Scale Weather System, Monthly Weather Review, 133, 1938–1960, https://doi.org/10.1175/MWR2958.1, 2005.

Caron, J.-F., Milewski, T., Buehner, M., Fillion, L., Reszka, M., Macpherson, S., and St-James, J.: Implementation of Deterministic Weather Forecasting Systems Based on Ensemble–Variational Data Assimilation at Environment Canada. Part II: The Regional System, Monthly Weather Review, 143, 2560–2580, https://doi.org/10.1175/MWR-D-14-00353.1, 2015.

Caron, J.-F., A., Z., and Co-authors: Regional Deterministic Prediction System (RDPS) – Update from version 4.2.0 to version 5.0.0, Technical note, Canadian Meteorological Center, Environment and Climate Change Canada, https://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/lib/technote_rdps-500_20160907_e.pdf, 2016.

Carrera, M. L., Bélair, S., Fortin, V., Bilodeau, B., Charpentier, D., and Doré, I.: Evaluation of Snowpack Simulations over the Canadian Rockies with an Experimental Hydrometeorological Modeling System, Journal of Hydrometeorology, 11, 1123–1140, https://doi.org/10.1175/2010JHM1274.1, 2010.

Carrera, M. L., Bélair, S., and Bilodeau, B.: The Canadian Land Data Assimilation System (CaLDAS): Description and Synthetic Evaluation Study, Journal of Hydrometeorology, 16, 1293–1314, https://doi.org/10.1175/JHM-D-14-0089.1, 2015.

Chikhar, K. and Gauthier, P.: Impact of Analyses on the Dynamical Balance of Global and Limited-Area Atmospheric Models: Impact of Analyses on the Dynamical Balance of Atmospheric Models, Quarterly Journal of the Royal Meteorological Society, 140, 2535–2545, https://doi.org/10.1002/qj.2319, 2014.

Côté, J., Desmarais, J.-G., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The Operational CMC–MRB Global Environmental Multiscale (GEM) Model. Part II: Results, Monthly Weather Review, 126, 1397–1418, https://doi.org/10.1175/1520-0493(1998)126<1397:TOCMGE>2.0.CO;2, 1998a.

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The Operational CMC–MRB Global Environmental Multiscale (GEM) Model. Part I: Design Considerations and Formulation, Monthly Weather Review, 126, 1373–1395, https://doi.org/10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2, 1998b.

Deacu, D., Fortin, V., Klyszejko, E., Spence, C., and Blanken, P.: Predicting the net basin supply to the Great Lakes with a hydrometeorological model, Journal of Hydrometeorology, 13, 1739–1759, https://doi.org/10.1175/JHM-D-11-0151.1, 2012.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Fairbairn, D., Barbu, A. L., Napoly, A., Albergel, C., Mahfouf, J.-F., and Calvet, J.-C.: The Effect of Satellite-Derived Surface Soil Moisture and Leaf Area Index Land Data Assimilation on Streamflow Simulations over France, Hydrol. Earth Syst. Sci., 21, 2015–2033, https://doi.org/10.5194/hess-21-2015-2017, 2017.

© Author(s) 2021. CC BY 4.0 License.

Fillion, L., Mitchell, H. L., Ritchie, H., and Staniforth, A.: The Impact of a Digital Filter Finalization Technique in a Global Data Assimilation System, Tellus A, 47, 304–323, https://doi.org/10/bbvz39, 1995.

Fletcher, S.: Data Assimilation for the Geosciences, Elsevier, 1 edn., 2017.

Fortin, V. and Gronewold, A. D.: Water Balance of the Laurentian Great Lakes, in: Encyclopedia of Lakes and Reservoirs, pp. 864–869, Springer, Dordrecht, https://doi.org/10.1007/978-1-4020-4410-6_268, 2012.

Fortin, V., Roy, G., Donaldson, N., and Mahidjiba, A.: Assimilation of Radar Quantitative Precipitation Estimations in the Canadian Precipitation Analysis (CaPA), Journal of Hydrology, 531, 296–307, https://doi.org/10.1016/j.jhydrol.2015.08.003, 2015.

Fortin, V., Roy, G., Stadnyk, T., Koenig, K., Gasset, N., and Mahidjiba, A.: Ten Years of Science Based on the Canadian Precipitation Analysis: A CaPA System Overview and Literature Review, Atmosphere - Ocean, 56, 178–196, https://doi.org/10.1080/07055900.2018.1474728, 2018.

Fry, L. M., Hunter, T. S., Phanikumar, M. S., Fortin, V., and Gronewold, A. D.: Identifying Streamgage Networks for Maximizing the Effectiveness of Regional Water Balance Modeling, Water Resources Research, 49, 2689–2700, https://doi.org/10.1002/wrcr.20233, 2013.

Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., Tegtmeier, S., Chabrillat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., and Zou, C.-Z.: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and Overview of the Reanalysis Systems, Atmospheric Chemistry and Physics, 17, 1417–1452, https://doi.org/10.5194/acp-17-1417-2017, 2017.

Gaborit, É., Fortin, V., Xu, X., Seglenieks, F., Tolson, B., Fry, L. M., Hunter, T., Anctil, F., and Gronewold, A. D.: A Hydrological Prediction System Based on the SVS Land-Surface Scheme: Efficient Calibration of GEM-Hydro for Streamflow Simulation over the Lake Ontario Basin, Hydrol. Earth Syst. Sci., 21, 4825–4839, https://doi.org/10.5194/hess-21-4825-2017, 2017.

Gagnon, N. and Co-authors: Global Ensemble Prediction System (GEPS) – Update from Version 4.0.1 to Version 4.1.1, Technical note, Canadian Meteorological Center, Environment and Climate Change Canada, https://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/lib/technote_geps-411_20151215_e.pdf, 2015.

Giorgi, F.: Thirty Years of Regional Climate Modeling: Where Are We and Where Are We Going next?, Journal of Geophysical Research: Atmospheres, 124, 5696–5723, https://doi.org/10.1029/2018JD030094, 2019.

Girard, C., Plante, A., Desgagné, M., McTaggart-Cowan, R., Côté, J., Charron, M., Gravel, S., Lee, V., Patoine, A., Qaddouri, A., Roch, M., Spacek, L., Tanguay, M., Vaillancourt, P. A., and Zadra, A.: Staggered Vertical Discretization of the Canadian Environmental Multiscale (GEM) Model Using a Coordinate of the Log-Hydrostatic-Pressure Type, Monthly Weather Review, 142, 1183–1196, https://doi.org/10/gcw8wg, 2013.

Girard, C., Plante, A., Desgagné, M., Mctaggart-Cowan, R., Côté, J., Charron, M., Gravel, S., Lee, V., Patoine, A., Qaddouri, A., Roch, M., Spacek, L., Tanguay, M., Vaillancourt, P., and Zadra, A.: Staggered vertical discretization of the canadian environmental multiscale (GEM) model using a coordinate of the log-hydrostatic-pressure type, Monthly Weather Review, 142, 1183–1196, https://doi.org/10.1175/MWR-D-13-00255.1, 2014.

Gronewold, A. D. and Rood, R. B.: Recent water level changes across Earth's largest lake system and implications for future variability, Journal of Great Lakes Research, 45, 1 – 3, https://doi.org/https://doi.org/10.1016/j.jglr.2018.10.012, 2019.

Gronewold, A. D., Fortin, V., Caldwell, R., and Noel, J.: Resolving Hydrometeorological Data Discontinuities along an International Border, Bulletin of the American Meteorological Society, https://doi.org/10.1175/BAMS-D-16-0060.1, 2017.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 Global Reanalysis, Quarterly Journal of the Royal Meteorological Society, pp. 1–51, https://doi.org/10.1002/qj.3803, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803, 2020.

Houtekamer, P. L., Deng, X., Mitchell, H. L., Baek, S.-J., and Gagnon, N.: Higher Resolution in an Operational Ensemble Kalman Filter, Monthly Weather Review, 142, 1143–1162, https://doi.org/10.1175/MWR-D-13-00138.1, 2013.

Lavaysse, C., Carrera, M., Bélair, S., Gagnon, N., Frenette, R., Charron, M., and Yau, M. K.: Impact of Surface Parameter Uncertainties within the Canadian Regional Ensemble Prediction System, Monthly Weather Review, 141, 1506–1526, https://doi.org/10.1175/MWR-D-11-00354.1, 2012.

Lespinas, F., Fortin, V., Roy, G., Rasmussen, P., and Stadnyk, T.: Performance Evaluation of the Canadian Precipitation Analysis (CaPA), Journal of Hydrometeorology, 16, 2045–2064, https://doi.org/10.1175/JHM-D-14-0191.1, 2015.

Li, X., Charron, M., Spacek, L., and Candille, G.: A Regional Ensemble Prediction System Based on Moist Targeted Singular Vectors and Stochastic Parameter Perturbations, Monthly Weather Review, 136, 443–462, https://doi.org/10.1175/2007MWR2109.1, 2008.

Lin, H., Gagnon, N., Beauregard, S., Muncaster, R., Markovic, M., Denis, B., and Charron, M.: GEPS-Based Monthly Prediction at the Canadian Meteorological Centre, Monthly Weather Review, 144, 4867–4883, https://doi.org/10.1175/MWR-D-16-0138.1, 2016.

Lin, Y. and Mitchell, K.: The NCEP Stage II/IV hourly precipitation analyses: development and applications, in: Preprints of the 19th Conference on Hydrology, American Meteorological Society, San Diego, CA, 9-13 January 2005, p. Paper 1.2, https://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/refs/stage2-4.19hydro.pdf, 2005.

Lott, N., Baldwin, R., and Jones, P.: The FCC Integrated Surface Hourly Database, A New Resource of Global Climate Data, Tech. Rep. 2001-01, U.S. National Climate Data Center, https://rda.ucar.edu/datasets/ds463.3/docs/ish-tech-report.pdf, 2001.

Lucas-Picher, P., Boberg, F., Christensen, J. H., and Berg, P.: Dynamical Downscaling with Reinitializations: A Method to Generate Finescale Climate Datasets Suitable for Impact Studies, Journal of Hydrometeorology, 14, 1159–1174, https://doi.org/10/f5fb5t, 2013.

Mahfouf, J.-F., Brasnett, B., and Gagnon, S.: A Canadian Precipitation Analysis (CaPA) Project: Description and Preliminary Results, Atmosphere-Ocean, 45, 1–17, https://doi.org/10.3137/ao.v450101, 2007.

Mai, J., Kornelsen, K. C., Tolson, B. A., Fortin, V., Gasset, N., Bouhemhem, D., Schäfer, D., Leahy, M., Anctil, F., and Coulibaly, P.: The Canadian Surface Prediction Archive (CaSPAr): A Platform to Enhance Environmental Modeling in Canada and Globally, Bulletin of the American Meteorological Society, 101, E341 – E356, https://doi.org/10.1175/BAMS-D-19-0143.1, 2020a.

Mai, J., Tolson, B., Shen, H., Gaborit, E., Fortin, V., Gasset, N., Awoye, H., Stadnyk, T., Fry, L., Bradley, E., Seglenieks, F., Temgoua, A., Princz, D., Gharari, S., Haghnegahdar, A., Elshamy, M., Razavi, S., Gauch, M., Lin, J., Ni, X., Yuan, Y., McLeod, M., Basu, N., Kumar, R., Rakovec, O., Samaniego, L., Attinger, S., Shrestha, N., Daggupati, P., Roy, T., Wi, S., Hunter, T., and Craig, J.: The Great Lakes Runoff Intercomparison Project Phase 3: Lake Erie (GRIP-E), Journal of Hydrologic Engineering, to be submitted, 2020b.

Mai, J., Tolson, B. A., Shen, H., Étienne Gaborit, Fortin, V., Gasset, N., Awoye, H., Stadnyk, T. A., Fry, L. M., Bradley, E. A., Seglenieks, F., Temgoua, A. G., Princz, D. G., Gharari, S., Haghnegahdar, A., Elshamy, M. E., Razavi, S., Gauch, M., Lin, J., Ni, X., Yuan, Y., McLeod, M., Basu, N., Kumar, R., Rakovec, O., Samaniego, L., Attinger, S., Shrestha, N. K., Daggupati, P., Roy, T., Wi, S., Hunter, T., Craig, J. R., and Pietroniro, A.: The Great Lakes Runoff Intercomparison Project Phase 3: Lake Erie (GRIP-E), Journal of Hydrologic Engineering (under review), 2021.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Mailhot, J., Bélair, S., Lefaivre, L., Bilodeau, B., Desgagné, M., Girard, C., Glazer, A., Leduc, A., Méthot, A., Patoine, A., Plante, A., Rahill, A., Robinson, T., Talbot, D., Tremblay, A., Vaillancourt, P., Zadra, A., and Qaddouri, A.: The 15-km version of the Canadian regional forecast system, Atmosphere-Ocean, 44, 133–149, https://doi.org/10.3137/ao.440202, 2006.

Marke, T., Mauser, W., Pfeiffer, A., and Zängl, G.: A Pragmatic Approach for the Downscaling and Bias Correction of Regional Climate Simulations: Evaluation in Hydrological Modeling, Geoscientific Model Development, 4, 759–770, https://doi.org/10.5194/gmd-4-759-2011, 2011.

McTaggart-Cowan, R. and Zadra, A.: Representing Richardson Number Hysteresis in the NWP Boundary Layer, Monthly Weather Review, 143, 1232–1258, https://doi.org/10.1175/MWR-D-14-00179.1, 2014.

McTaggart-Cowan, R., Girard, C., Plante, A., and Desgagné, M.: The Utility of Upper-Boundary Nesting in NWP, Monthly Weather Review, 139, 2117–2144, https://doi.org/10/bjz8gq, 2011.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., and Shi, W.: North American Regional Reanalysis, Bulletin of the American Meteorological Society, 87, 343–360, https://doi.org/10.1175/BAMS-87-3-343, 2006.

Muñoz Sabater, J., Dutra, E., Schepers, D., Albergel, C., Boussetta, S., Agusti-Panareda, A., Zsoter, E., and Hersbach, H.: ERA5-Land: An improved version of the ERA5 reanalysis land component, p. 20, Joint International Surface Working Group and Satellite Applications Facility on Land Surface Analysis Workshop, IPMA, Lisbon, Portugal, 2018.

Noilhan, J. and Mahfouf, J. F.: The ISBA Land Surface Parameterisation Scheme, Global and Planetary Change, 13, 145–159, https://doi.org/10.1016/0921-8181(95)00043-7, 1996.

Noilhan, J. and Planton, S.: A Simple Parameterization of Land Surface Processes for Meteorological Models, Monthly Weather Review, 117, 536–549, https://doi.org/10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2, 1989.

Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Liu, Q., Mahanama, S. P. P., and Touré, A.: Assessment and Enhancement of MERRA Land Surface Hydrology Estimates, Journal of Climate, 24, 6322–6338, https://doi.org/10.1175/JCLI-D-10-05033.1, 2011.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, Journal of Climate, 24, 3624–3648, https://doi.org/10.1175/JCLI-D-11-00015.1, 2011.

Smith, G. C., Roy, F., Mann, P., Dupont, F., Brasnett, B., Lemieux, J.-F., Laroche, S., and Bélair, S.: A New Atmospheric Dataset for Forcing Ice-Ocean Models: Evaluation of Reforecasts Using the Canadian Global Deterministic Prediction System: CGRF Dataset for Forcing Ice-Ocean Models, Quarterly Journal of the Royal Meteorological Society, 140, 881–894, https://doi.org/10.1002/qj.2194, 2014.

Takacs, L. L., Suárez, M. J., and Todling, R.: Maintaining Atmospheric Mass and Water Balance in Reanalyses, Quarterly Journal of the Royal Meteorological Society, 142, 1565–1573, 2016.

Wang, X. L., Xu, H., Qian, B., Feng, Y., and Mekis, E.: Adjusted Daily Rainfall and Snowfall Data for Canada, Atmosphere-Ocean, 55, 155–168, https://doi.org/10.1080/07055900.2017.1342163, 2017.

Yoshimura, K. and Kanamitsu, M.: Dynamical Global Downscaling of Global Reanalysis, Monthly Weather Review, 136, 2983–2998, https://doi.org/10.1175/2008MWR2281.1, 2008.

Zadra, A., Gauthier, J.-P., and Leroux, A.: GenPhysX: A user's guide to input/output and methods, Tech. rep., Canadian Meteorological Centre, Environment Canada, 2008.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

List of Figures

1 Scheme illustrating the main components of the method followed to produce a land surface and precipitation reanalysis. . . . 33
2 Flowchart of the method followed to produce a land surface and precipitation reanalysis. . . . 34
3 Flowchart illustrating the approach to build 24-h accumulation and produce 24-h precipitation analysis at 12 UTC based on 6-h to 18-h leadtime RDRS outputs from integrations started at 00 UTC and 12 UTC. . . . 35
4 Comparison of atmospheric results from operational RDPS (black dashed line), RDRS+Open-Loop (blue dashed-dotted line), and RDRS+CaLDAS (orange continuous line) after 24-h integration against radiosonde for the whole North America (from left to right: wind modulus, geopotential height, absolute temperature, dew point depression): a) February/March 2011; b) July/August 2011. In each of the sub-figures the curves on the left show the bias while the curves on the right show the STDE. . . . 36
5 Comparison of reforecast surface results from RDRS+Cycle (blue) and RDRS+CaLDAS (orange) against SYNOP observations in terms of STDE (top curves, solid lines) and bias (bottom curves, dashed lines) computed based on monthly intervals for the whole North America and the 2013-2014 years: a) absolute temperature; b) dew point temperature. Results based on a combination of 6-h to 17-h leadtime from integrations starting at 00 UTC and 12 UTC. . . . 37
6 Time series of RMSE (against SYNOP observations) from Operational RDPS (black dashed line), RDRS-10 (orange continuous line) and RDRS-15 (blue dashed-dotted line) computed based on monthly intervals for the whole North America and the 2011-2017 years: a) absolute temperature; b) dew point temperature; c) wind speed. Results only based on a combination of 6-h to 17-h leadtime from integrations starting at 0 UTC and 12 UTC. . . . 38
7 Seasonal diurnal cycles of surface results from operational RDPS (black with square markers) and RDRS-10 (orange with round markers) against SYNOP observations in terms of STDE (top curves, solid lines) and bias (bottom curves, dashed lines) for entire North America over the period 2011-2017 (from left to right: winter (DJF); spring (MAM); summer (JJA); autumn (SON)): a) absolute temperature; b) dew point temperature. Results only based on the 6-h to 17-h leadtime from integrations starting at 00 UTC (filled markers) and 12 UTC (empty markers). . . . 39
8 Precipitation accumulation per seasons [mm/month] over years 2010-2014 for : a) CaPA-24h based on RDRS-10; b) CaPA-24h based on RDRS-15; c) NOAA Multisensor Precipitation Estimates (Stage IV) 24-h analysis; d) Global Precipitation Climatology Project (GPCP) monthly analysis . . . 40
9 Comparison of CaPA-24h background field (dashed lines) and analyses (continuous lines) at 10 km (orange lines) and 15 km (blue lines) against manual synoptic stations over years 2010-2014. Each row corresponds to a season (winter, spring, summer and fall), each column to a score (probability of detection, false alarm ratio, frequency bias index, equitable threat score and partial mean). . . . 41
10 Comparison to observed Lake Erie inflows (grey dots) of GEM-Hydro simulations based on RDRS-10 + CaPA-24h (continuous orange line), RDRS-15 + CaPA-24h (dashed-dotted blue line) and RDPS + CaPA-24h (dashed black line) forcing for years 2011-2014: (a) daily inflows [m^3 s^{-1}]; (b) ranked daily inflows, log scale [m^3 s^{-1}] (c) cumulative inflow over the five-year period [m]. . . . 42

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Figures

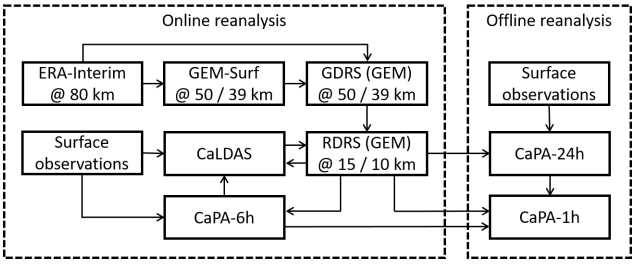


Figure 1. Scheme illustrating the main components of the method followed to produce a land surface and precipitation reanalysis.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

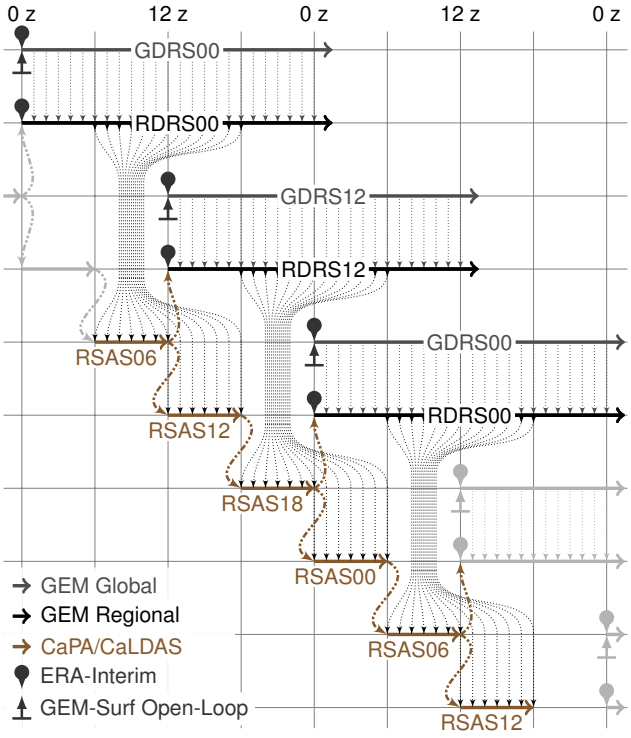


Figure 2. Flowchart of the method followed to produce a land surface and precipitation reanalysis.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

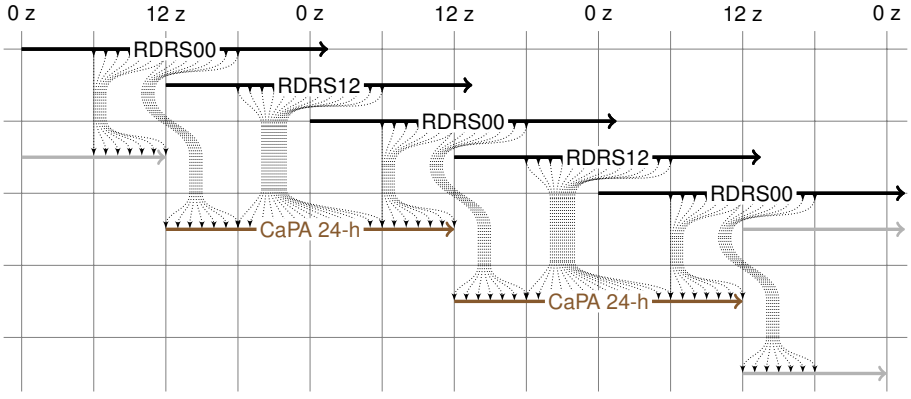


Figure 3. Flowchart illustrating the approach to build 24-h accumulation and produce 24-h precipitation analysis at 12 UTC based on 6-h to 18-h leadtime RDRS outputs from integrations started at 00 UTC and 12 UTC.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

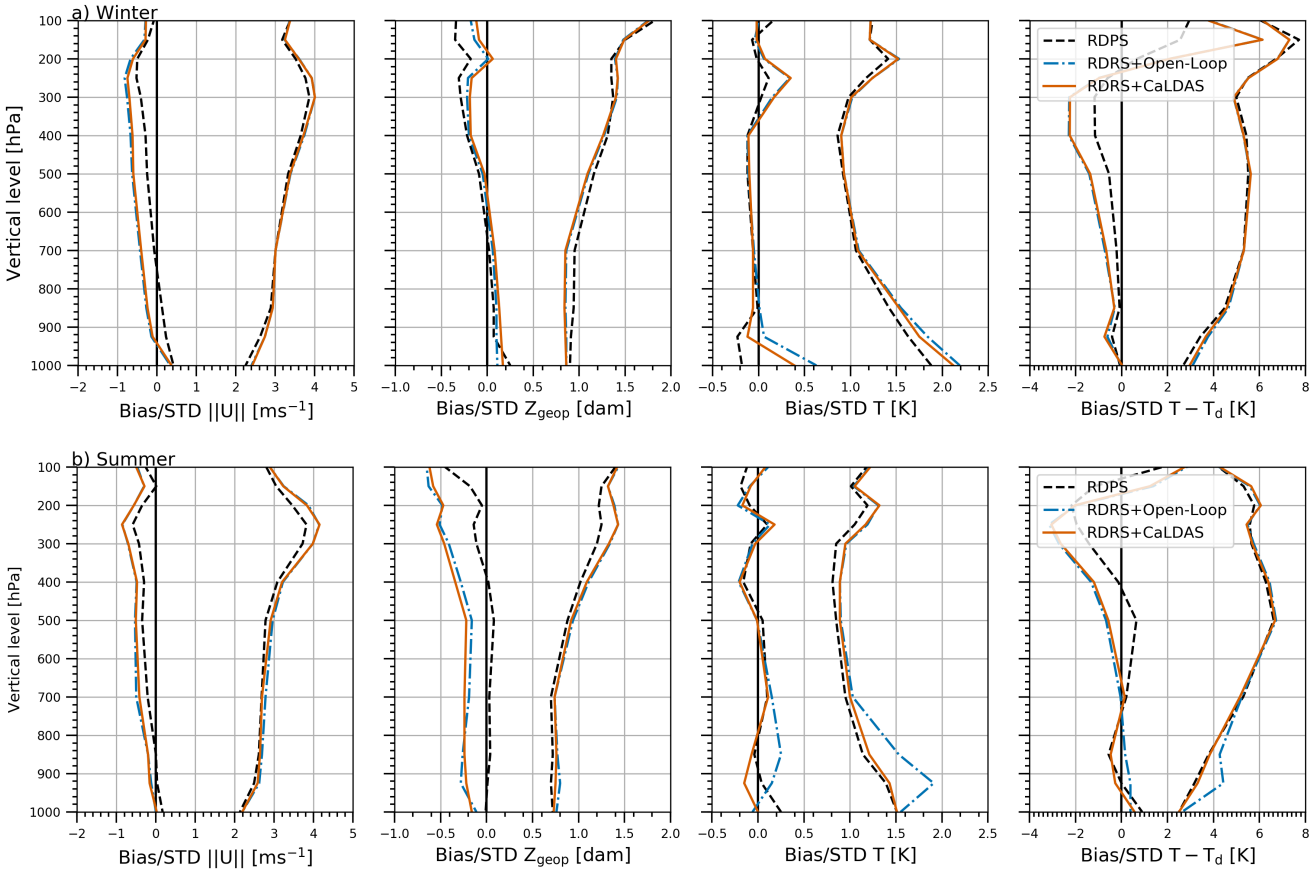


Figure 4. Comparison of atmospheric results from operational RDPS (black dashed line), RDRS+Open-Loop (blue dashed-dotted line), and RDRS+CaLDAS (orange continuous line) after 24-h integration against radiosonde for the whole North America (from left to right: wind modulus, geopotential height, absolute temperature, dew point depression): a) February/March 2011; b) July/August 2011. In each of the sub-figures the curves on the left show the bias while the curves on the right show the STDE.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

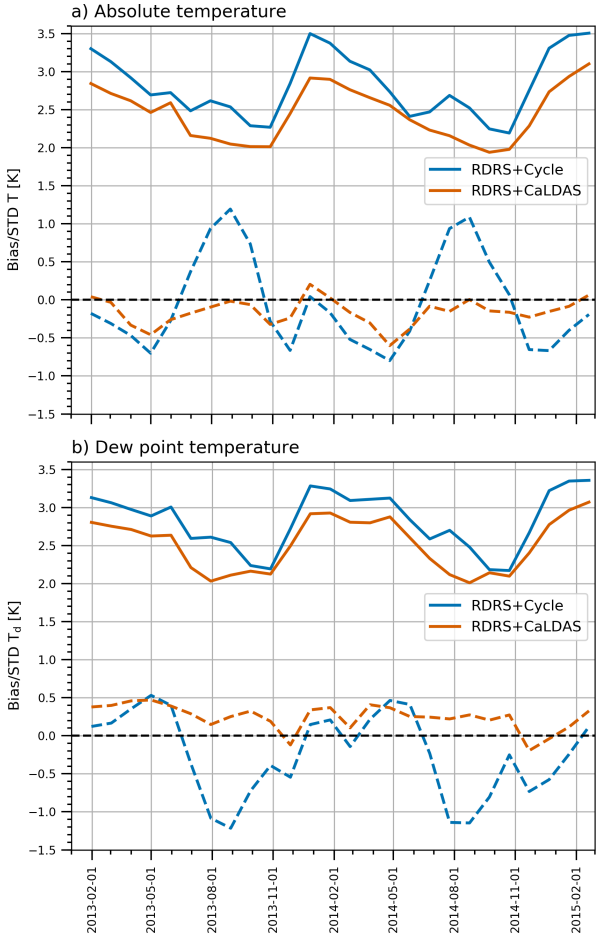


Figure 5. Comparison of reforecast surface results from RDRS+Cycle (blue) and RDRS+CaLDAS (orange) against SYNOP observations in terms of STDE (top curves, solid lines) and bias (bottom curves, dashed lines) computed based on monthly intervals for the whole North America and the 2013-2014 years: a) absolute temperature; b) dew point temperature. Results based on a combination of 6-h to 17-h leadtime from integrations starting at 00 UTC and 12 UTC.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

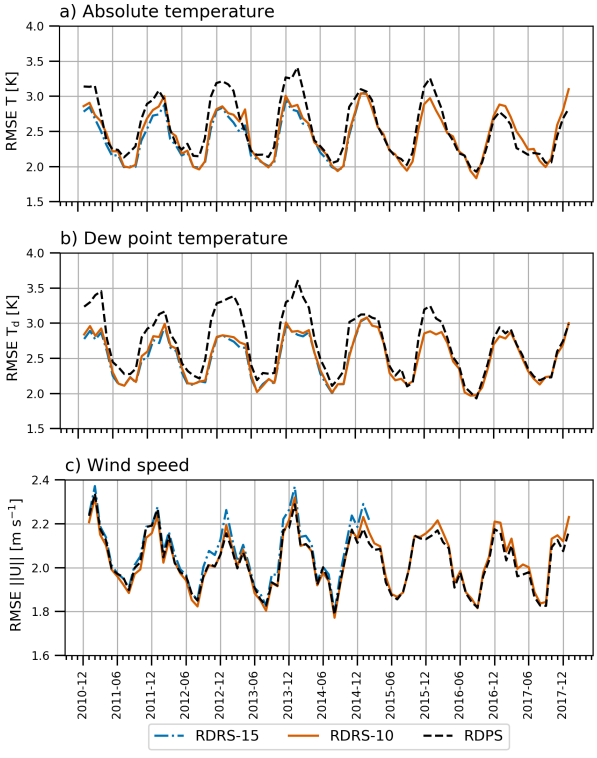
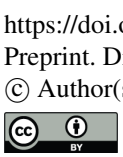


Figure 6. Time series of RMSE (against SYNOP observations) from Operational RDPS (black dashed line), RDRS-10 (orange continuous line) and RDRS-15 (blue dashed-dotted line) computed based on monthly intervals for the whole North America and the 2011-2017 years: a) absolute temperature; b) dew point temperature; c) wind speed. Results only based on a combination of 6-h to 17-h leadtime from integrations starting at 0 UTC and 12 UTC.

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.



Hydrology and Earth System Sciences
Discussions

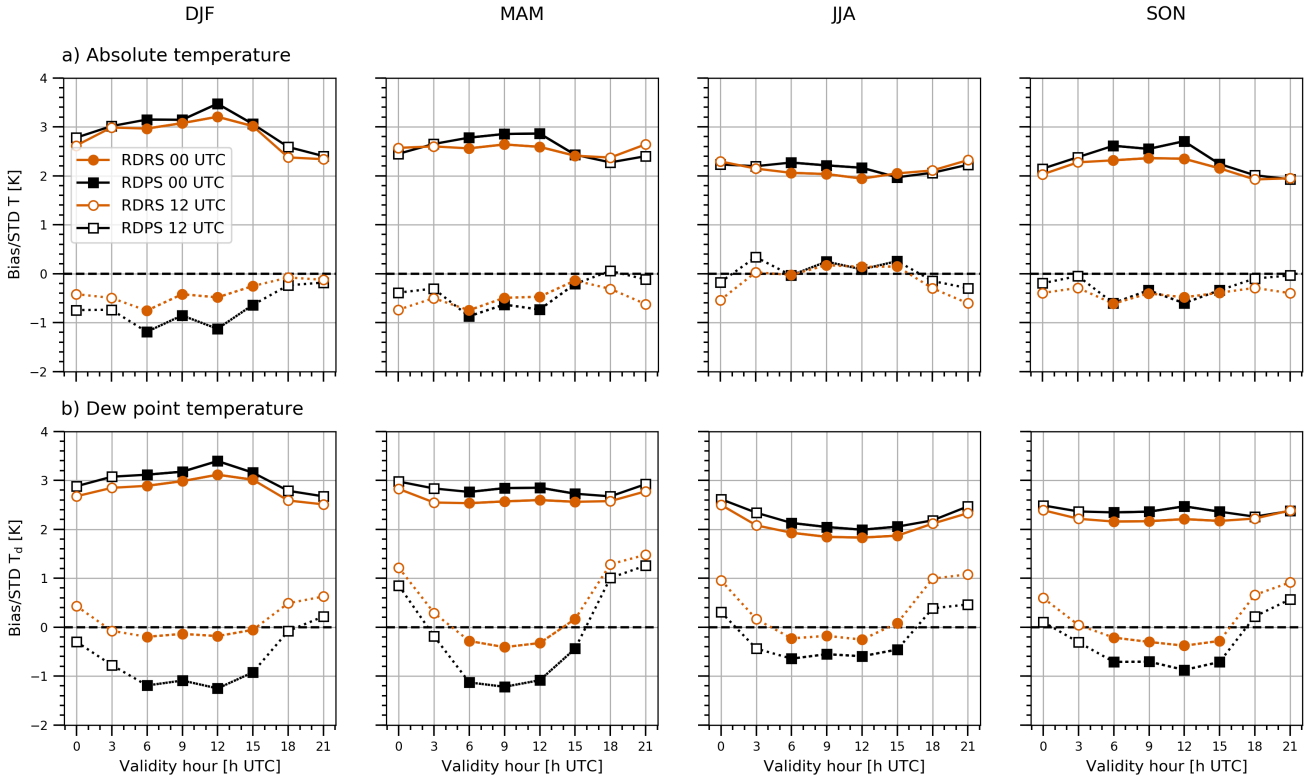


Figure 7. Seasonal diurnal cycles of surface results from operational RDPS (black with square markers) and RDRS-10 (orange with round markers) against SYNOP observations in terms of STDE (top curves, solid lines) and bias (bottom curves, dashed lines) for entire North America over the period 2011-2017 (from left to right: winter (DJF); spring (MAM); summer (JJA); autumn (SON)): a) absolute temperature; b) dew point temperature. Results only based on the 6-h to 17-h leadtime from integrations starting at 00 UTC (filled markers) and 12 UTC (empty markers).

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

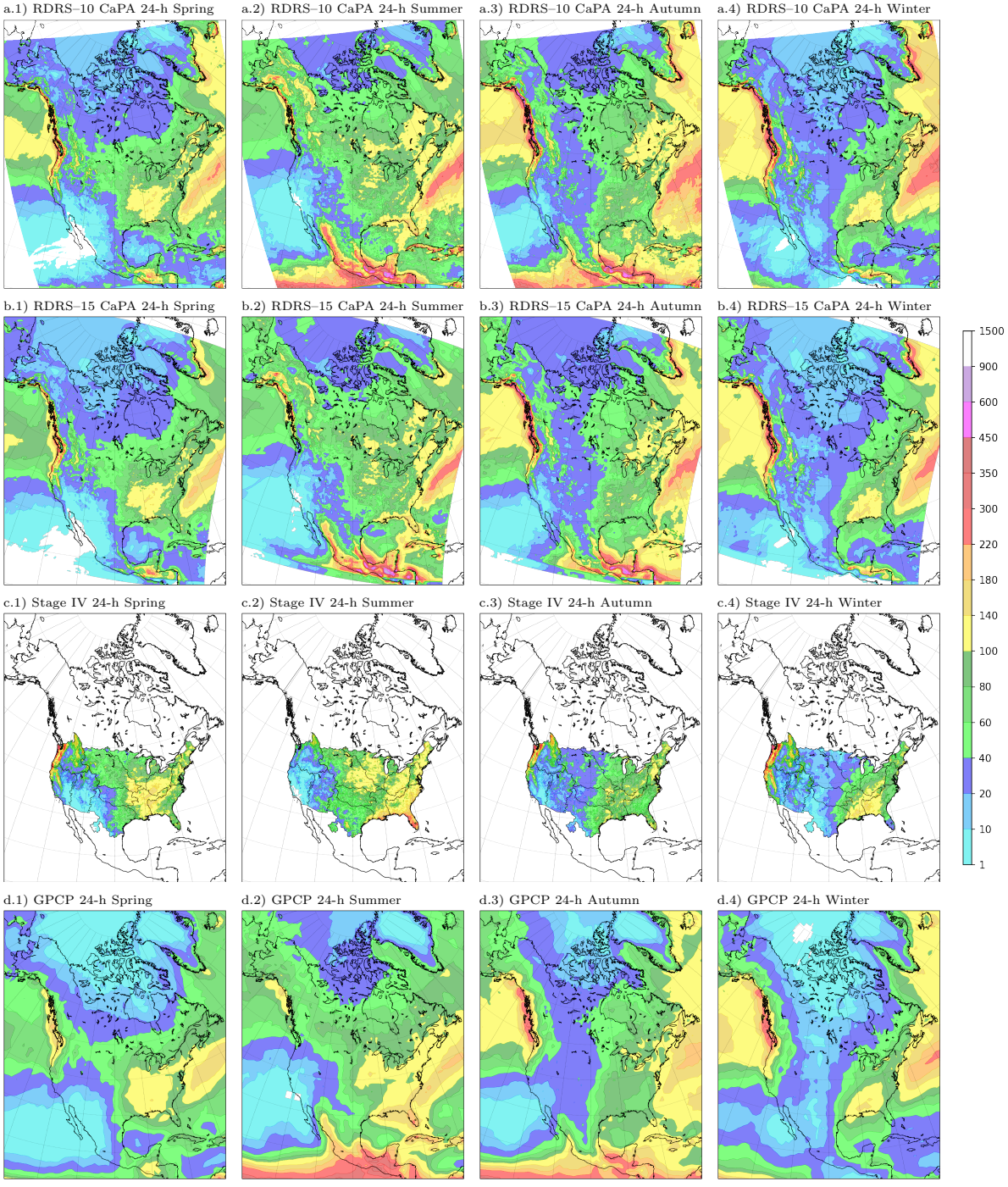


Figure 8. Precipitation accumulation per seasons [mm/month] over years 2010-2014 for : a) CaPA-24h based on RDRS-10; b) CaPA-24h based on RDRS-15; c) NOAA Multisensor Precipitation Estimates (Stage IV) 24-h analysis; d) Global Precipitation Climatology Project (GPCP) monthly analysis

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

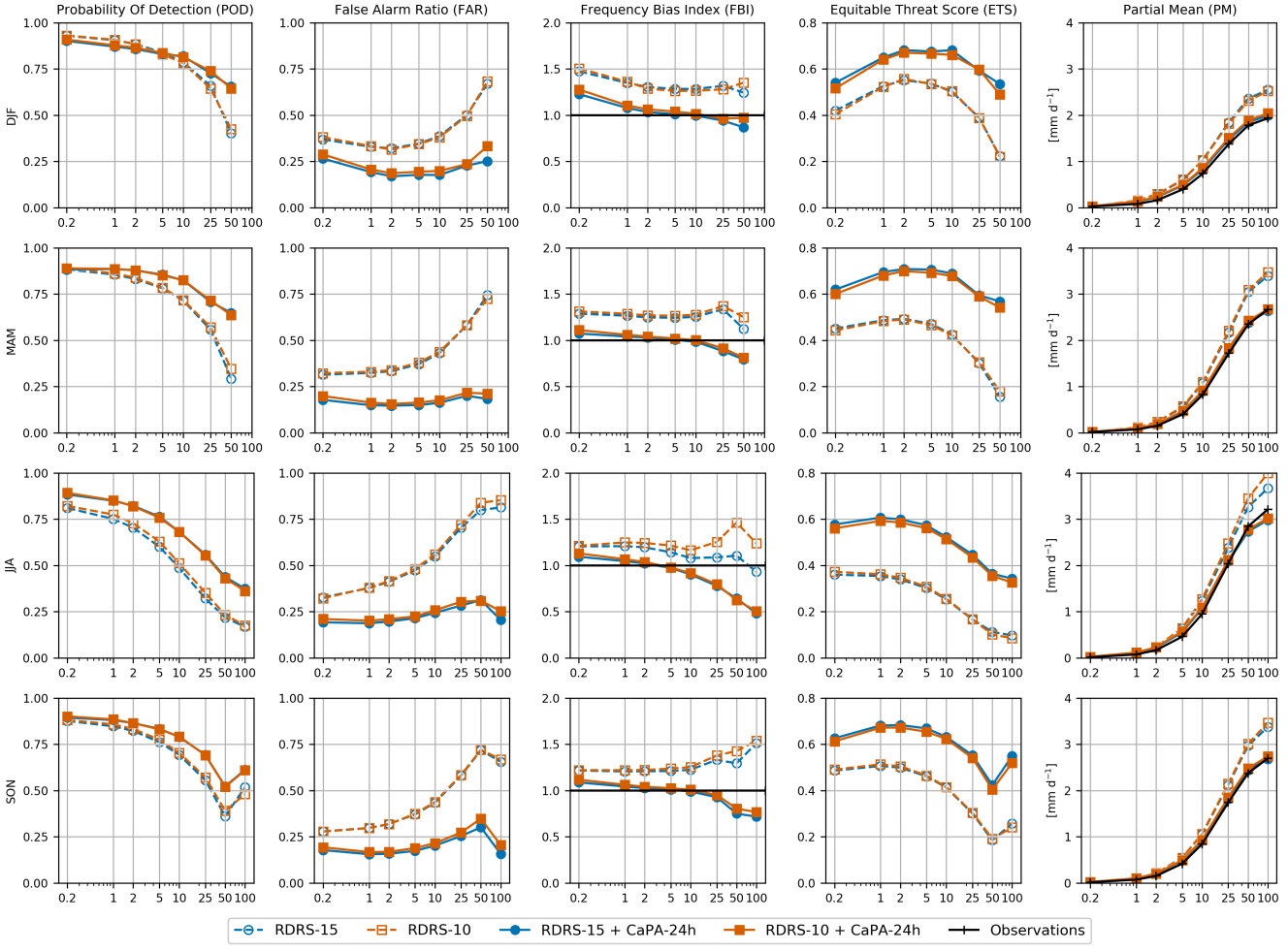


Figure 9. Comparison of CaPA-24h background field (dashed lines) and analyses (continuous lines) at 10 km (orange lines) and 15 km (blue lines) against manual synoptic stations over years 2010-2014. Each row corresponds to a season (winter, spring, summer and fall), each column to a score (probability of detection, false alarm ratio, frequency bias index, equitable threat score and partial mean).

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

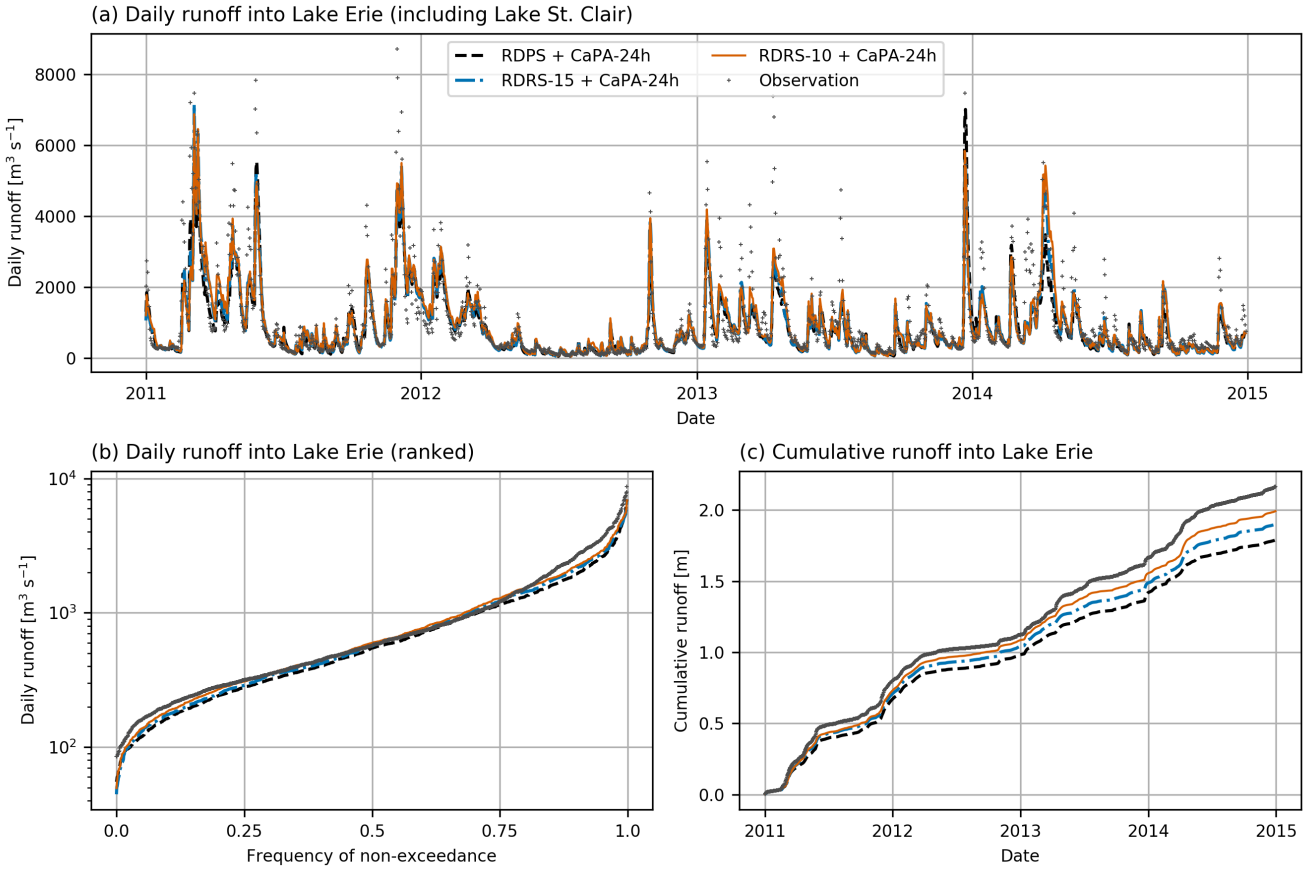


Figure 10. Comparison to observed Lake Erie inflows (grey dots) of GEM-Hydro simulations based on RDRS-10 + CaPA-24h (continuous orange line), RDRS-15 + CaPA-24h (dashed-dotted blue line) and RDPS + CaPA-24h (dashed black line) forcing for years 2011-2014: (a) daily inflows [$\mathrm{m}^3\ \mathrm{s}^{-1}$]; (b) ranked daily inflows, log scale [$\mathrm{m}^3\ \mathrm{s}^{-1}$] (c) cumulative inflow over the five-year period [m].

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

List of Tables

1 Dynamic kernel and physical parametrizations common to the GDRS and RDRS. 45
2 Main differences in the configuration of the GDRS and RDRS. 46
3 Surface networks and variables used by CaLDAS and CaPA. 47
4 Summary of average changes (in %) of RMSE (against SYNOP observations) between RDRS+Open-Loop and RDRS+CaLDAS for February/March 2011 and July/August 2011. For each season, two columns show results from 24-h integrations starting at either 00 UTC or 12 UTC. Changes of more than 10% are shown in bold, negative values are underlined. 48
5 Summary by seasons of average changes (in %) of RMSE (against SYNOP observations) between operational RDPS, RDRS-10 and RDRS-15 for 2011-2014 years. For each season, results are based on a combination of 6-h to 17-h leadtime from integrations starting at 0 UTC and 12 UTC. Changes of more than 10% are shown in bold, negative values are underlined. 49
6 Bias, STDE, RMSE and MAE of snow density, snow water equivalent and snow depth computed from January 1, 2014 to January 31, 2015 with and without the fix for the bug identified in the RDRS (and in the operational RDPS) when estimating the maximum value for snow density as a function of snow depth. Best score for each snow characteristic is shown in bold. Number of observations used for the evaluation reported in bottom two rows. 50

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Tables

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Table 1. Dynamic kernel and physical parametrizations common to the GDRS and RDRS.

Model version	Global Environmental Multiscale (GEM) model version 4.8-LTS
Formulation	Hydrostatic primitive equations
Numerical technique	Finite differences: Arakawa C grid in the horizontal and Charney-Phillips grid in the vertical Girard et al. (2013)
Time integration	Implicit, semi-Lagrangian (3-D), 2 time-level Côté et al. (1998b, a)
Independent variables	x, y, η and time
Prognostic variables	E-W and N-S winds, temperature, specific humidity and logarithm of surface pressure, liquid water content, Turbulent kinetic energy (TKE)
Initialization	Diabatic Digital Filter Fillion et al. (1995)
Radiation	Solar and infrared using a correlated-k distribution (CKD) (Li and Barker, 2005)
Surface scheme	Interaction Soil-Biosphere-Atmosphere (ISBA) surface scheme: Mosaic approach with 4 types: land, water, sea ice and glacier Bélair et al. (2003a, b)
Deep convection	Kain & Fritsch scheme (Kain and Fritsch, 1990 and 1993)
Shallow convection	Girard scheme calculates turbulent fluxes in partially saturated air (see in Mailhot et al. 1998). Kuo Transient scheme (see Bélair et al., 2005)
Stable precipitation	Sundqvist et al., 1989 scheme (see Pudykiewicz et al., 1992). For QPF evaluations, see Bélair et al. (2009)
Gravity wave drag	Orographic: McFarlane (1987) and McFarlane et al. (1987) with coefficient 8.E-6. Non-orographic: Hines (1997a,b)
Low level (orographic) blocking	Parameterized (Lott and Miller, 1997; Zadra et al., 2003) with coefficient 1.0
Turbulent mixing (vertical diffusion)	Turbulent kinetic energy (Benoît et al., 1989; Bélair et al., 1999 for RDRS-15 and Bélair et al., 2005 for RDRS-10) except near the surface and in the upper-troposphere. Prandtl number taken to be 1.0. Includes Richardson number hysteresis McTaggart-Cowan and Zadra (2014)

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table 2. Main differences in the configuration of the GDRS and RDRS.

Prediction system	GDRS–50	GDRS–39	RDRS–15	RDRS–10
Grid	800×400 uniform	Yin-Yang uniform	726×556 rotated lat-lon	960×1080 rotated lat-lon
Horizontal resolution	0.45° ($\sim$50 km)	0.35° ($\sim$39 km)	0.1375° ($\sim$15 km)	0.09° ($\sim$10 km)
Vertical levels	45	45	48	80
Lid [hPa]	0.1	0.1	10	0.1
Lid piloting	No	No	Yes	No
Time step [s]	900	900	450	300
Surface analysis	GEM-Surf forced by ERA-Interim	GEM-Surf forced by ERA-Interim	CaLDAS	CaLDAS
Time period covered	2010–2014	1980–2018	2010–2014	2000–2017 (1980–2018 planned)

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table 3. Surface networks and variables used by CaLDAS and CaPA.

Network	**Domain covered**	**Availability at CCMEP**	**Variables used by CaLDAS**	**Variables used by CaPA-6h**	**Variables used by CaPA-24h**
METAR	North America	1992-present	T, T_d	P, T, $\|\|U\|\|$	P, T, $\|\|U\|\|$
SWOB	North America	2013-present	T, T_d, S_d	Not used	P, T, $\|\|U\|\|$
SYNOP	North America	1992-present	T, T_d, S_d	P, T, $\|\|U\|\|$	P, T, $\|\|U\|\|$
AdjDlyRS	Canada	1980-present	Not used	Not used	P
RMCQ	Province of Quebec	2011-present	Not used	Not used	P, T, $\|\|U\|\|$
SHEF	USA	1998-present	Not used	Not used	P, T, $\|\|U\|\|$

T: temperature, T_d: dew point temperature, S_d: snow depth, P: total precipitation, $||U||$: wind speed

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table 4. Summary of average changes (in %) of RMSE (against SYNOP observations) between RDRS+Open-Loop and RDRS+CaLDAS for February/March 2011 and July/August 2011. For each season, two columns show results from 24-h integrations starting at either 00 UTC or 12 UTC. Changes of more than 10% are shown in bold, negative values are underlined.

RMSE change (%) RDRS+Open-Loop vs RDRS+CaLDAS		**2011-02-01 to 2011-03-31**		**2011-07-01 to 2011-08-31**	
		00z	**12z**	**00z**	**12z**
Alaska plus Canadian Arctic	T_d	7.45	8.87	6.07	5.22
	T	5.59	7.63	8.09	8.11
	$\|\|U\|\|$	-0.10	-0.12	0.17	-0.01
Canada West	T_d	5.52	4.32	**15.51**	7.48
	T	3.76	2.77	9.97	**13.88**
	$\|\|U\|\|$	-0.13	-0.11	0.91	2.05
Canada East	T_d	2.41	2.76	6.22	5.75
	T	2.20	4.22	3.77	3.96
	$\|\|U\|\|$	0.13	0.10	-0.12	-0.23
United States of America West	T_d	**15.86**	**14.24**	**22.35**	**21.32**
	T	5.84	5.20	**10.33**	**10.66**
	$\|\|U\|\|$	0.42	1.02	1.23	2.41
United States of America East	T_d	8.35	9.47	**21.29**	**23.13**
	T	-3.61	0.55	**16.86**	**17.77**
	$\|\|U\|\|$	0.16	0.70	0.86	3.42
Canada	T_d	5.29	5.22	**12.68**	7.67
	T	4.30	5.04	8.19	**10.72**
	$\|\|U\|\|$	-0.02	-0.03	0.40	0.89
United States of America	T_d	**12.83**	**12.23**	**22.10**	**21.85**
	T	2.63	3.40	**13.69**	**13.86**
	$\|\|U\|\|$	0.32	0.85	1.10	2.98
North America	T_d	6.88	6.82	**15.20**	**11.73**
	T	3.60	4.45	9.34	**11.30**
	$\|\|U\|\|$	0.00	0.06	0.49	1.22

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table 5. Summary by seasons of average changes (in %) of RMSE (against SYNOP observations) between operational RDPS, RDRS-10 and RDRS-15 for 2011-2014 years. For each season, results are based on a combination of 6-h to 17-h leadtime from integrations starting at 0 UTC and 12 UTC. Changes of more than 10% are shown in bold, negative values are underlined.

RMSE change (%) 2010-12-01 to 2015-02-28		RDPS vs RDRS-10				RDRS-15 vs RDRS-10			
		DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
Alaska plus Canadian Arctic	T_d	**10.45**	**13.43**	**22.06**	**15.20**	-2.04	-1.31	-0.87	-1.04
	T	**10.83**	**11.40**	2.34	**13.81**	-1.70	-1.91	-2.52	-1.30
	$\|\|U\|\|$	-0.09	0.24	1.90	0.95	0.78	0.15	0.33	1.13
Canada West	T_d	**12.78**	**14.71**	6.76	7.60	-2.37	-0.85	0.96	-1.15
	T	1.66	4.06	3.04	3.78	-5.16	-2.74	0.49	-2.07
	$\|\|U\|\|$	-0.47	0.97	0.84	0.73	0.40	0.52	1.34	1.12
Canada East	T_d	3.87	5.41	7.41	5.25	2.04	-1.56	-0.70	2.20
	T	3.68	**-12.29**	-4.98	4.10	5.71	-6.40	-2.30	2.21
	$\|\|U\|\|$	-1.24	-1.05	0.52	0.07	4.43	2.49	1.99	4.13
USA West	T_d	**21.08**	**13.50**	4.53	**15.05**	-3.56	-2.16	-0.61	-2.76
	T	**24.79**	**10.73**	9.79	**21.10**	-3.31	-3.90	-2.43	-1.47
	$\|\|U\|\|$	2.00	1.52	1.35	1.84	1.89	0.06	0.64	1.24
USA East	T_d	3.99	3.49	0.79	3.35	-2.48	-1.70	-0.49	-1.96
	T	8.31	5.24	6.59	9.06	-6.08	-2.68	-0.90	-3.05
	$\|\|U\|\|$	0.28	0.59	1.39	1.27	1.49	2.23	1.92	3.04
Canada	T_d	8.57	**11.18**	9.58	7.69	-1.03	-1.13	0.36	-0.33
	T	2.29	-1.28	0.01	5.07	-1.07	-4.04	-1.22	-0.84
	$\|\|U\|\|$	-0.91	-0.08	0.95	0.53	2.16	1.35	1.54	2.46
USA	T_d	**13.47**	**10.00**	3.12	**10.74**	-2.87	-1.94	-0.58	-2.40
	T	**17.35**	8.70	8.95	**16.89**	-4.81	-3.28	-1.77	-2.18
	$\|\|U\|\|$	0.92	0.90	1.20	1.48	1.60	1.29	1.44	2.27
Mexico	T_d	**17.36**	**19.30**	**20.81**	**24.42**	1.70	0.79	1.78	1.48
	T	**11.65**	5.42	7.21	**11.11**	-2.90	-3.77	-2.19	-2.06
	$\|\|U\|\|$	-0.10	-0.70	-1.35	0.10	2.00	0.77	1.10	1.40
North America	T_d	**10.51**	**11.15**	7.70	9.03	-1.26	-1.22	0.03	-0.71
	T	8.03	2.67	2.68	9.27	-1.78	-3.91	-1.24	-1.10
	$\|\|U\|\|$	-0.35	0.29	0.98	0.61	2.23	1.41	1.45	2.56

https://doi.org/10.5194/hess-2021-41
Preprint. Discussion started: 9 March 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Table 6. Bias, STDE, RMSE and MAE of snow density, snow water equivalent and snow depth computed from January 1, 2014 to January 31, 2015 with and without the fix for the bug identified in the RDRS (and in the operational RDPS) when estimating the maximum value for snow density as a function of snow depth. Best score for each snow characteristic is shown in bold. Number of observations used for the evaluation reported in bottom two rows.

		Bias		**STDE**		**RMSE**		**MAE**		**Number of obs.**	
Snow characteristic	**Unit**	w/bug	w/fix	w/bug	w/fix	w/bug	w/fix	w/bug	w/fix	w/bug	w/fix
Snow Density	$\mathrm{kg\ m^{-3}}$	144	**7**	90	**79**	169	**78**	155	**57**	2779	2732
Snow depth	cm	-12	**-10**	18	**16**	22	**19**	14	**12**	4399	4388
SWE	$\mathrm{kg\ m^{-2}}$	**11**	-20	67	**43**	68	**54**	52	**34**	4349	4338