



Implementing a parsimonious variable contributing area algorithm for the prairie pothole region in the HYPE modelling framework

Mohamed Ismaiel Ahmed^{a,b,*}, Kevin Shook^c, Alain Pietroniro^d, Tricia Stadnyk^a, John W. Pomeroy^c, Charlotta Pers^e, David Gustafsson^e

^a Department of Geography, University of Calgary, Calgary, AB, T2N 1N4, Canada

^b Irrigation and Hydraulic Department, Faculty of Engineering, Cairo University, Cairo University Rd, Giza, Egypt

^c Centre for Hydrology, University of Saskatchewan, Saskatoon, SK, S7N 1K2, Canada

^d Schulich School of Engineering, University of Calgary, Calgary, AB, T2N 4V8, Canada

^e Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

ARTICLE INFO

Handling Editor: Daniel P. Ames

Keywords:

HYPE
Contributing area
Prairie pothole
Streamflow
Hysteresis

ABSTRACT

The North American prairie region is known for its poorly defined drainage system with numerous surface depressions that lead to variable contributing areas for streamflow generation. Current approaches of representing surface depressions are either simplistic or computationally demanding. In this study, a variable contributing area algorithm is implemented in the Hydrological Predictions for the Environment (HYPE) model and evaluated in the Canadian prairies. HYPE's local lake module is replaced with a Hysteretic Depressional Storage (HDS) algorithm to estimate the variable contributing fractions of subbasins. The modified model shows significant improvements in simulating the streamflows of two prairie basins in Saskatchewan, Canada. The modified model can replicate the hysteretic relationships between the water volume and contributing area of the basins. With the inclusion of the HDS algorithm in HYPE, the global HYPE modelling community can now simulate an important hydrological phenomenon, previously unavailable in the model.

1. Introduction

The North American prairie region is characterized by a semi-arid climate and flat topography with millions of land/surface depressions from glacial origin known as prairie potholes (Pomeroy et al., 2007; Zhang et al., 2009). These potholes can retain significant amounts of surface runoff (Shook and Pomeroy, 2011; van der Kamp and Hayashi, 2009) leading to hysteretic relationships between the contributing area and storage within the land depressions (Shook et al., 2013) and a non-linear response of the basin (Ahmed et al., 2021). The potholes within the basin have differing areas, volumes, available storage, and connections with surrounding potholes, which can cause the amount of retained water to vary greatly across the landscape (Ahmed et al., 2020a; Shook et al., 2021). During dry conditions, these depressions are not typically connected to the stream network, and they form internally drained basins wherein they do not contribute to streamflow (Godwin and Martin, 1975; Hayashi et al., 2003; Martin, 2001). However, during wet conditions, the ponds in the depressions can reach their depressions' capacity, their surface area expands, and they become connected

(merge) with surrounding ponds to contribute flow to the river network (Ahmed et al., 2021; Shook and Pomeroy, 2011). Connections between the depressions can be established by both surface and subsurface flow (Ahmed et al., 2021; Ameli and Creed, 2017). This connection is known as the fill and spill mechanism, wherein a depression starts to spill water to the surrounding/downstream areas and/or depressions after being filled (Shaw et al., 2012), and was first identified by Spence and Woo (2003) for lakes in the Canadian shield. This mechanism leads to contributing areas that vary in time and space based on the prior states of depressional storage and the magnitude of the event (Ahmed et al., 2021). Furthermore, the region is dominated by cold region processes, such as snow accumulation, redistribution, and ablation (Fang and Pomeroy, 2007), and freezing/thawing of the soil and limited infiltrability of frozen soil (Gray and Landine, 1988; Pomeroy et al., 2007), which when added to the land depressions complexities makes it particularly challenging to simulate the hydrology of that region. Therefore, the prairie region has been known as "the graveyard of hydrological models" wherein most traditional models fail (Ahmed et al., 2020b; Mekonnen et al., 2014, 2016).

* Corresponding author. Department of Geography, University of Calgary, Calgary, AB, T2N 1N4, Canada.

E-mail address: mohamedismaiel.ahmed@ucalgary.ca (M.I. Ahmed).

<https://doi.org/10.1016/j.envsoft.2023.105769>

Received 14 October 2022; Received in revised form 1 March 2023; Accepted 3 July 2023

Available online 4 July 2023

1364-8152/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Several conceptual approaches have been proposed to represent the land depressions in hydrologic models (Ahmed et al., 2020b; Evenson et al., 2016; Mekonnen et al., 2014, 2016; Muhammad et al., 2019; Zeng et al., 2020). A fixed multiplier approach was proposed for the HYPE modelling framework, wherein the non-effective area of the basin (identified from the historic Prairie Farm Rehabilitation (PFRA) maps; Agriculture and Agri-Food Canada, 2013) is multiplied by a fixed coefficient to represent the non-contributing fraction of the basin (Stadnyk et al., 2020). This simplified approach assumes that the contributing area is temporally fixed (non-variable), which has been shown to be an invalid assumption for the prairie region (Mekonnen et al., 2014). The Probability Distribution Model (PDM; Moore, 2007) concept has been adapted and implemented in many hydrological models in the prairie region such as: HYPR (Ahmed et al., 2020b); SWAT (Mekonnen et al., 2016), and MESH (Mekonnen et al., 2014). The PDM concept simulates all the depressions in the basin as a conceptual lumped storage unit and it uses a probability distribution (Pareto or Power type) to represent the variable contributing area. Other conceptual attempts have been made to represent the depressions within the basin as distinct series of cascading storage units. These approaches were adapted for multiple versions of the SWAT model (Muhammad et al., 2019; Nasab et al., 2017), wherein each depression contributes flow to the downstream areas/depressions after reaching their capacity. The above-mentioned conceptual approaches may not be adequate to represent the complex dynamics/connection of the prairie depressions or the variable contributing area. Their inability to replicate the hysteretic relationships between the contributing area and depressional storage is problematic (Ahmed et al., 2021).

More explicit/physical approaches have been developed to simulate the actual spatiotemporal dynamics of prairie depressions. Simple hydraulic models (e.g., PRIMA, Ahmed et al., 2020a; WDPM, Shook et al., 2013; Shook and Pomeroy, 2011) and a more sophisticated hydrodynamic model (e.g., FLUXOS-OVERFLOW; Costa et al., 2020) were developed as fully distributed and explicit solutions to simulate the prairie depressional complexities. The models simulate the movement of water over the landscape and quantify the spatial change in depressional connectivity and storage. FLUXOS-OVERFLOW is more complex and much more computationally demanding compared to WDPM and PRIMA. Unlike WDPM, PRIMA calculates the travel time of water over prairie landscape, giving it the potential to be implemented in hydrological models of that region. The PRIMA model was coupled with the MESH land surface model (MESH-PRIMA; Ahmed et al., 2021) to simulate the prairie hydrology more accurately. The modified MESH model (MESH-PRIMA) was able to provide more reliable streamflow simulation when compared against more conceptual approaches (e.g., the PDM concept). More importantly, the MESH-PRIMA model was the first and only model that coupled a hydrological model with a distributed representation of the depressions to simulate the spatiotemporal changes in the contributing area over the tested basin (Smith Creek Research Basin in SK, Canada) and was able to simulate the hysteretic relationships between contributing area and storage. However, these physically based approaches are computationally and data demanding, which limits their general applicability to medium sized basins.

The Pothole Cascade Model (PCM; Pomeroy et al., 2014; Shook et al., 2013) was developed to simplify the WDPM hysteretic behavior using a representative statistical sample of depressions. It is numerically simple and was implemented in the Cold Regions Hydrological Model (CRHM; Pomeroy et al., 2014) and was found to provide simulations that matched known hysteretic relationships between depressional storage, contributing area and discharge efficiently. This permitted accurate simulations of the influence of wetlands and variable contributing area in the Smith Creek Research Basin from an uncalibrated hydrological model (Pomeroy et al., 2014).

There is a need for a more parsimonious approach that maintains a full representation of the depressional complexities, and is more computationally efficient, than detailed physically based approaches;

and is more applicable in large-scale domains. It is important to have such a model to better understand both floods and droughts in these landscapes. This is especially evident after the 2011, 2013, and 2014 flooding events and the droughts of 2000 and the 2020s in the Canadian prairies. This region is expected to experience more of these extreme events under a changing climate and the potholes are crucial to the sustained economic development in the region and are also a critical part of the natural ecosystem.

The main objective of this paper is to provide a solution to the pothole and variable contributing area representation problem in large-scale or global hydrologic models. This is accomplished through modifications of the HYdrological Predictions for the Environment (HYPE; Lindström et al., 2010) model. In this paper, HYPE is modified by replacing its pre-built-in lake (local/internal lake) module with a Hysteretic Depressional Storage and contributing area functions (HDS) to be able to simulate the variable contributing area and runoff of prairie basins. The modified model is referred to as HYPE-HDS. It is hypothesized that HYPE-HDS will improve HYPE's performance in replicating the complex variable contributing area, streamflow, and hydrology of the North American prairie environment. This will be useful not only for North America, but also for the Siberian and Pan-Arctic areas, that have similar hydrological controls as the prairie landscape, giving this approach global applicability. We introduce an improved and integrated solution to the hydrological simulations worldwide (e.g., WW-HYPE, Arheimer et al., 2020).

2. Methodology

In order to accommodate the dynamics of variable contributing areas, the HYPE model was modified by replacing its local lake module with the Hysteretic Depressional Storage (HDS) algorithm to improve its ability to simulate prairie streamflow. The subsequent changes to the model were then evaluated in the streamflow simulation of HYPE-HDS as compared to the classic version of HYPE using its current lake module. Two well-studied basins in the Canadian prairies were used to evaluate and test the model and to show the potential improvements to the model by introducing the HDS algorithm to HYPE. The basin storage and contributing area curves are analyzed for HYPE-HDS to better understand and quantify how the model simulated prairie basins and the connection between the depressions. The contributing area and the contributing fraction of the basin mean the same thing in this study and are used interchangeably. A detailed methodology outlining the basin, the model changes and the HDS algorithm is described below.

2.1. HYPE model

2.1.1. Overview

The HYdrological Predictions for the Environment (HYPE, Lindström et al., 2010) model was developed by the Swedish Meteorological and Hydrological Institute (SMHI) to facilitate the simulation of various hydrological processes across the different landscapes. HYPE was based initially on the HBV model (Lindström et al., 1997), but it includes additional empirical and physical methods to estimate land-surface's water and energy fluxes and can also preserve the spatial heterogeneity across the landscape. HYPE is computationally efficient compared to more sophisticated counterparts, making it more applicable to large-scale watersheds and more suitable for operational purposes (Lindström et al., 2010). HYPE shows potential to simulate the streamflow and various hydrological signatures across various scales in Canada and the entire globe (Ahmed et al., 2023; Arheimer et al., 2020; Bajracharya et al., 2020; Stadnyk et al., 2020; Tefs et al., 2021).

HYPE is a flexible and dynamic semi distributed hydrological model that can simulate hydrological fluxes and substance transport over the landscapes (Fig. 1), including but not limited to the following: snow and glacier processes (e.g., precipitation phase change, snow accumulation and ablation, glacier volume changes and melt), soil water plant

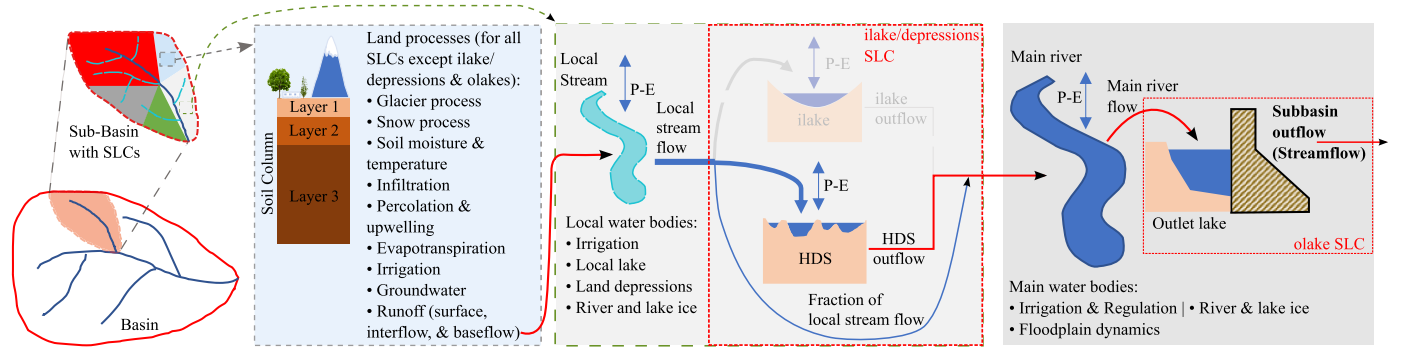


Fig. 1. Schematic diagram of the typical hierarchical modelling structure for each subbasin within the HYPE modelling framework with the inclusion of HDS. The *ilake* module (faded out component) was replaced by HDS to simulate the pothole complexities within the subbasin.

relationship (frozen soils dynamics, evaporation, infiltration, percolation, macropore flow), lateral runoff generation, river and lake processes (temperature, ice formation, outflow rating curves), wetland dynamics, sediment and substances (e.g. nitrogen and phosphorus) transport through the landscape and water bodies, water management processes (controlled lakes/dams, irrigation withdrawal, water abstraction and transfer), deep processes (groundwater flow and aquifer dynamics), and tracers (Arheimer et al., 2020; Lindström et al., 2010). There are multiple algorithms/modules that can be used to calculate the same processes and the user needs to specify which algorithms are to be used in the model configuration. For example, snowmelt can be estimated using simplified methods (degree-day approach) or calculated using physical methods (energy-budget approach). Evapotranspiration can be calculated using one of the following methods: temperature dependent (a conceptual method that is a function of air temperature, temperature threshold, and evaporation rate; SMHI, 2016), modified Jensen-Haise/McGuinness, modified Hargreaves-Samani, Priestly-Taylor, and FAO Penman-Monteith.

HYPE is a vector-based system, which requires the hydrological information at the subbasin level identified through watershed delineation. Each subbasin needs to be discretized into Soil and Land Cover classes (SLC), which are similar to the concept of Hydrological Response Units (HRUs), and the soil column for each SLC is discretized into three layers, by default. The soil column can be further discretized into seven soil layers (in Northern latitudes) to better simulate freezing/thawing of the soil column and permafrost (Bajracharya et al., 2023). The hydrological calculations for fluxes and states are made at the SLC scale. Then, the fluxes from all SLCs in a subbasin are aggregated to provide averaged values. The model runs on a daily time resolution, by default, but can run on a sub-daily basis (currently with 1 hour as the smallest time step). It also requires a minimum of two input forcings: precipitation and mean temperature, but the forcing fields can include shortwave radiation, relative humidity, wind speed, and minimum and maximum air temperature depending on the algorithms invoked in the model configuration by the user.

2.1.2. HDS description

The Hysteretic Depressional Storage (HDS) algorithm is based on hysteretic functions between the variable contributing area of the subbasin and the depressional storage. These functions were described by Shook et al. (2021) and assume that the relationship between the contributing area and storage is linear for subbasins dominated by a large number of land depressions. This permits simulation of all depressions within the subbasin as a conceptual lumped storage unit with a hysteretic relationship between contributing area and storage; a key feature to successfully simulating prairie streamflow. Accordingly, HDS does not simulate the depressions as distinct (individual) depressions (spatially distributed units), rather, it simulates the effects of the integration of all depressions within a system and their connection at the

subbasin level as a lumped (integrated/lumped meta-depression) unit.

The HDS algorithm requires two inputs: the runoff depth from the upland area and the net input vertical fluxes depth ($NIVF = \text{precipitation} - \text{evaporation} - \text{infiltration/leakage}$) to the depressions. The workflow is based on a series of calculations at each time step to update the controlling state variables. First, HDS calculates the water areal fraction of the depressions (WF_t , which is the fraction of the subbasin that is covered by water) based on the Hayashi and Van Der Kamp (2000) expression, relating the area and volume of water stored in a depression, adjusted to simulate all the depressions in a subbasin:

$$WF_t = SC \times VF_t^{2/P} \times WF_{max} \quad (1)$$

where SC is a scaling constant (model parameter), VF_t is the volume fraction of the depressions for the current time step t , represented by the ratio between the current depressional storage depth (DS_t , m) and the maximum possible depressional storage depth (DS_{max} , m; model parameter), P is the power controlling the relationship between area and depth (model parameter), and WF_{max} is the maximum possible water areal fraction of the depressions (fraction of the subbasin that is covered by water when the depressions are full, model parameter). Second, the algorithm calculates the upland area fraction of the subbasin (UF_t , that is not covered by water) for the current time step as follows:

$$UF_t = 1 - WF_t \quad (2)$$

Third, the algorithm calculates the depressional storage depth (DS_t , m) for the current time step (t), which is a function of the depressional storage depth (DS_{t-1} , m) of previous time step ($t-1$) and the change in depressional storage Δd_t . DS_t is expressed as follows:

$$DS_t = DS_{t-1} + \Delta d_t \quad (3)$$

The change in depressional storage (Δd_t) is derived from the water budget expressed as follows:

$$\Delta d_t = NIVF_t \times WF_t + R_t \times UF_t \times (1 - CF_t) \quad (4)$$

where $NIVF_t$ is the net input vertical fluxes depth ($NIVF_t = \text{precipitation} - \text{evaporation} - \text{infiltration/leakage}$; m) to the depressions, R_t is the upland runoff depth (m), and CF_t is the variable contributing fraction of the subbasin. Fourth, the algorithm calculates the net depressional outflow depth (DO_t , m) that reaches the river network for the current time step (t) as:

$$DO_t = \begin{cases} (DS_t - DS_{max}) \times WF_t + R_t \times UF_t \times CF_t, & DS_t > DS_{max} \\ R_t \times UF_t \times CF_t, & DS_t \leq DS_{max} \end{cases} \quad (5)$$

When $DS_t > DS_{max}$, the DS_t is set to equal DS_{max} after calculating DO_t . It is noteworthy that the HDS algorithm does not use storage-discharge rating curves to adjust DO_t values, since they are not applicable to the numerous land depressions within any typical prairie basin because it is nearly impossible to fit one curve to all depressions or to represent the

integrated effect of a system of depressions.

Lastly, the algorithm calculates the variable contributing area/fraction (CF_{t+1}) to be used by the model for the next time step ($t+1$) using the following equation:

$$CF_{t+1} = \begin{cases} CF_t, & \Delta d_t = 0 \\ 0, & \Delta d_t < 0 \\ CF_t + \Delta VF_t \times S_t, & \Delta d_t > 0 \end{cases} \quad (6)$$

where ΔVF_t is the change in the volume fraction occurred due to adding Δd_t ($\Delta VF_t = \Delta d_t / DS_{max}$), and S_t is the slope of the line connecting the current point (state as defined by VF and CF) to the (1,1) point (maximum volume fraction and maximum contributing fraction). S_t can be expressed as follows:

$$S_t = \max\left(\frac{1 - CF_t}{1 - VF_t}, 0\right) \quad (7)$$

The HDS algorithm requires four parameters: DS_{max} , WF_{max} , SC , and P based on the landscape configuration of the depressions. The first two parameters can be obtained by identifying the depressions (sinks) from the terrain DEM using GIS analysis. For example, the “DepthInSink” function from the Whitebox GIS (Lindsay, 2016) can be used to generate a depressional depth raster. The mean value of the depressional depth raster can be calculated to obtain the DS_{max} value. WF_{max} can be estimated as the number of cells containing values (in the depressional depth raster) divided by the total number of cells of the DEM. Alternatively, WF_{max} can be estimated from available landcover data by estimating the areal fraction of the water class (identified from landcover data) within the study domain. The remaining two parameters (SC and P) can be calibrated or set to recommended values (based on DS_{max} and WF_{max}) from the literature (Hayashi and Van Der Kamp, 2000). It is important to distinguish between water fraction (WF_t) and contributing fraction (CF_t). WF_t represents the fraction of the subbasin that is covered by water (lake area), while CF_t represents the fraction of the subbasin that is capable of contributing flow/runoff to the stream network. The former has implications for the state of the depressional storage and evapotranspiration, while the latter affects the quantity of runoff that translates into streamflow. It is also important to distinguish upland area (UF_t) fraction from contributing area fraction (CF_t). UF_t is the lumped/integrated catchment area of the depressions, which is the area that drains water into the depressions and not to the stream network. Conversely, CF_t is the area that contributes (drain) flow to the subbasin outlet or main river. Contributing area can include the upland area (UF_t) and the water fraction (WF_t) if the depressions are well connected and can translate water to the stream network.

2.1.3. HDS implementation in HYPE

In this study, the existing lake (ilake) module of HYPE, which handles all local/internal lakes within the subbasin as a lumped lake, was replaced with the HDS algorithm to allow HYPE to simulate aspects of the variable contributing area of the subbasin (Fig. 1). By default, any lake component of a hydrological model (like the ilake component of HYPE) requires the runoff depth and net vertical flux as inputs to the lake system. In the HYPE modelling framework, HYPE calculates the upland runoff (surface and subsurface water) from all SLCs within the subbasin (except the ilake/depressions SLC) and vertical net input (precipitation – evaporation) and pass them to the HDS algorithm (similar to how HYPE handles inputs/outputs of the regular ilake, Fig. 1). Then, the HDS algorithm quantifies the changes in the depressional storage, contributing area, and net outflow using the equations shown above. The net outflow is then passed back to HYPE to be moved downstream and used by other processes (Fig. 1). HDS is applied to each subbasin that has a value for the SLC fraction of depressions, and HDS parameters and states differ for each simulated subbasin.

The main difference between ilake and the HDS algorithm is the way runoff is generated. ilake generates runoff only when water exceeds a

specific threshold using a simple rating curve equation and a contributing area provided as a fixed fraction (typically 100%) of the subbasin area. However, the HDS algorithm calculates the runoff as a function of the variable contribution area, which is not related to a specific threshold and does not use any rating curve criteria for outflow calculations.

Some design choices (based on HYPE’s model structural configurations/limitations) were required to successfully implement HDS within HYPE, which can be summarized as follows:

1. WF_t was set to WF_{max} at all time steps because of the constraints of the HYPE model, which does not allow for variable lake/water area in the subbasin. This means that in this instance, Equation (1) was not used. In models that could accommodate variable lake area, Equation (1) should be included to make WF_t time varying and to have more physical representation of the process and more accurate estimation of the evaporative fluxes from the basin based on the variable wet fraction of the basin.
2. Only one parameter (DS_{max}) was added to the HYPE model, as a result of the modification, to be used by the HDS algorithm. SC and P parameters (used to update WF_t) were not included in the model because HYPE does not allow for variable lake area (point 1). The remaining parameter (WF_{max}) was identified from the depressions SLC defined through the model setup process.
3. Infiltration/leakage losses from the depressions were ignored because HYPE does not simulate such processes at the local water body (lakes/depressions) level and losses were limited to evaporation. This assumption can be considered valid for the semi-arid climate of the prairie region since most of the infiltrated water (leakage from the bottom of depressions) is being consumed by the growing vegetation on the moisture margin of the depression and then subsequently lost to evapotranspiration (Hayashi et al., 1998, 2003, 2016). Furthermore, much of the Canadian prairies are underlain by deep deposits of glacial till that restrict infiltration (Hayashi et al., 1998).
4. HDS does not include the crucial gatekeeping function of large depressions that was first noted by Phillips et al. (2011) in northern Canada and is important in some prairie basins (Pomeroy et al., 2010; Shook et al., 2021). HYPE has a module to simulate the gatekeeping effects (olake module), but it assumes that the large depression/lake will always be near the outlet of the basin, which may not always be true in the prairies. Therefore, we did not include the gatekeeping effects to test the applicability/performance of the hysteretic effects of the depressions on streamflow only (HDS algorithm) without complicating this experiment by including the gatekeeping effects.

2.2. Study area

The Smith Creek Research Basin (SCRB) and St. Denis National Wildlife Area above pond 90 (SDNWA-90) in Saskatchewan, Canada (Fig. 2) were selected to test the performance of HYPE-HDS because there is a good understanding of the variable contributing area across these complex landscapes through past modelling efforts (Ahmed et al., 2020a, 2021; Dumanski et al., 2015; Fang et al., 2010; Fang and Pomeroy, 2008; Mengistu et al., 2016; Shook et al., 2013, 2021; Shook and Pomeroy, 2011). These basins represent two extremes of variable contributing basins with land depressions within the prairie environment. SCRB is a highly cultivated basin with quite flat topography (mean slope of 3%) and a dominant land cover of cropland and pasture. SCRB has an area of 435 km² and includes more than 10,000 depressions with areas >100 m² (Fang et al., 2010). Accordingly, only 13% of the basin can contribute flow to the outlet for events with return period of 2 years or smaller according to the static PFRA non-effective area map (Fig. 2). SCRB has a well-defined stream with a clear valley near the outlet of the basin (Fig. 2). Shook et al. (2021) demonstrated that because of the

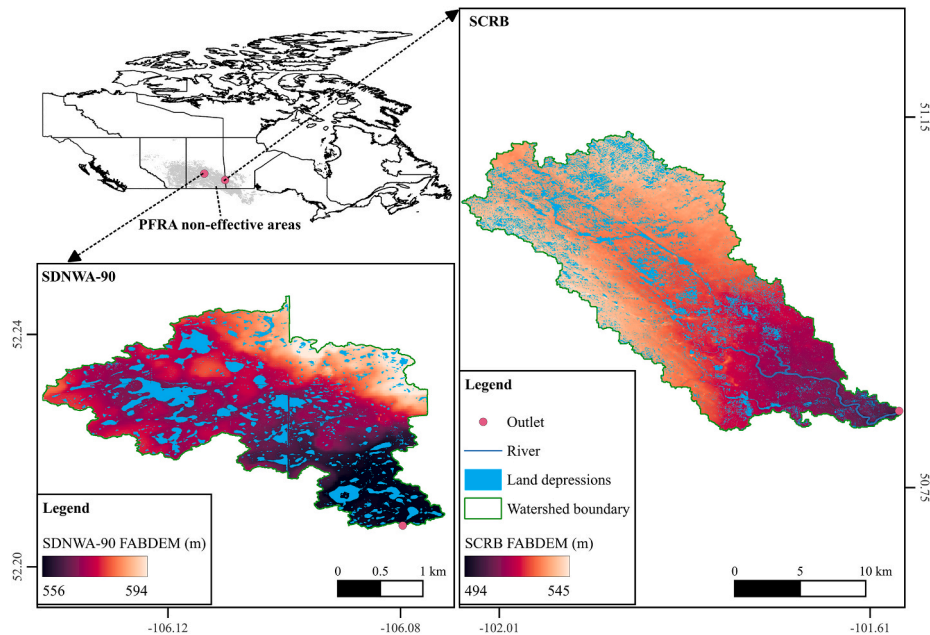


Fig. 2. A layout of the general extent of the non-effective area (identified by PFRA), the Smith Creek Research Basin (SCRB), and the St. Denis National Wildlife Area above pond 90 (SDNWA-90) test sites with land depressions. Elevation data for the basins are obtained from the FABDEM DEM. The vertical line in the SDNWA-90 DEM (center of the plot) represents a geographically prominent road that is detected in the used elevation data.

absence of relatively large depressions, SCRB was not greatly affected by gatekeeping.

Unlike SCRB, SDNWA-90 has a hummocky landscape with an average slope of approximately 12% with no clear or defined stream network or drainage system (Fig. 2). SDNWA-90 is also dominated by cropland cover and has an approximate area of 10 km². SDNWA-90 has more than 1000 land depressions having areas greater than 100 m² and is dominated by large depressions that are scattered over the landscape (Ahmed et al., 2020a, Fig. 2), which do cause gatekeeping.

2.3. Model setup and input datasets

Both HYPE-HDS and the original HYPE model with its local lake module (HYPE-ilake) were tested on SCRB and SDNWA-90 to show the improved prairie characterizations introduced by implementing the HDS algorithm in HYPE. Each study site was represented in HYPE using one subbasin with four SLCs (HRUs, computational units) and one special SLC to represent the depressions. The HYPE model setup has two subbasins in total; the same configuration is used by HYPE-ilake and HYPE-HDS. The different SLCs were identified based on a combination of the landcover and soil types within the basins. The land cover classes within the basins were identified from the North American Land Change Monitoring System product (NALCMS), which provides a consistent representation of the landcover for North America at a 30 m spatial resolution, based on the 2010 Landsat satellite imagery (Latifovic et al., 2016). The soil type/texture information were identified from the Global Soil Dataset for Earth System Modelling (GSDE), which includes various soil properties (e.g., particle-size, nutrients, organic matters, etc.) around the globe at ~ 1 km grid resolution (Shangguan et al., 2014). The Forest And Buildings removed Copernicus DEM (FABDEM; Hawker et al., 2022) was used to estimate the total depression depth for each study site that is used by both ilake (*lake_depth*) and HDS (*DS_{max}*) algorithms to maintain consistency across the model setups. FABDEM is a ~30 m global DEM with forest and building removed from the original Copernicus GLO 30 DEM.

The model forcing fields (precipitation and temperature) were obtained from the Regional Deterministic Reanalysis System (RDRS-v2; Gasset et al., 2021) dataset, which available at a 10 km resolution on an

hourly time scale from 2000 to 2017 over North America. RDRS-v2 is based on the reanalysis of the Global Environmental Multiscale atmospheric model (GEM; Mailhot et al., 2006) and the Canadian Precipitation Analysis product (CaPA; Lespinas et al., 2015) introduced to provide more accurate estimation of the meteorological fluxes over North America compared to the raw output of GEM-CaPA. The forcings were averaged over each basin and aggregated to a daily time step (total precipitation and minimum, maximum, and mean temperature). The streamflow records of SCRB were obtained from the Water Survey of Canada (WSC) from 1975 to 2017 for gauge 05ME007 at the outlet of the watershed. For SDNWA-90, the streamflow records were obtained from the Global Institute for Water Security (GIWS) at the University of Saskatchewan for the period of 2011–2014 (Ahmed et al., 2022).

A set of simplified and conceptual algorithms within the HYPE modelling framework calculated the fluxes for both HYPE-ilake and HYPE-HDS. For example, snowmelt was estimated using the degree-day approach, and evapotranspiration was calculated as a function of air temperature and the soil moisture storage. The model simulates the precipitation phase, snow accumulation, snow sublimation, snowmelt, frozen soil infiltration, freezing/thawing of the soil column, evapotranspiration, surface runoff, and flow routing. Blowing snow redistribution to depressions was ignored (Fang and Pomeroy, 2009; Pomeroy et al., 1993) because HYPE does not simulate such complex phenomena. Although these simplified approaches were proven to be inadequate in simulating the prairie snow dynamics (Gray and Landine, 1988; Pomeroy et al., 1993), they have been deployed with calibration for the purpose of simulating the streamflow of the basin (Ahmed et al., 2020b; Mekonnen et al., 2016; Muhammad et al., 2018; Zeng et al., 2020). This simplified form of the HYPE model (simplified flux calculations and one subbasin for each watershed) was used as an extreme case to test the applicability of the HDS algorithm in order to better assess the reliability of the improvements. More importantly, these algorithms are mostly/commonly used (and might be the only option) for areas with limited input data and forcings (such as the Siberian and Pan-arctic data). By demonstrating that the HDS algorithm can improve the streamflow simulations, even with simplified fluxes, it follows that it should also work with more physically representative and accurate flux simulations.

2.4. Simulation period and model calibration

The simulation periods were set to have both RDRS-v2 forcings and recorded flows. For SCRB, the simulation period was set from 2000 to 2017 (hydrological year, October to September). The period from Jan-2000 to Sep-2000 was used as a spin-up period. This period is believed to be sufficient for model initializing since it was a relatively dry period (Ahmed et al., 2020b; Ahmed et al., 2023; Bonsal et al., 2013) and the depressions were relatively empty. The period from 2001 to 2011 was considered to be the calibration period, while the period from 2012 to 2017 was used for model validation. The calibration period included the 2011 flooding event because it is known that flooding events are useful for model parameter identification, and it was shown to improve model performance over that region (Ahmed et al., 2020b, 2021). SDNWA-90 was used as an additional validation site. No calibration was done to fit the simulated flows of SDNWA-90 to observed values. The model was calibrated to fit the SCRB flows for the period 2001–2011 only and the parameters were transferred to SDNWA-90, except for the HDS parameter (DS_{max}) that was obtained from the terrain analysis of the FABDEM. The simulation period for SDNWA-90 was set from 2000 to 2017 (similar to SCRB as they are both used the same model setup). However, only the results of the period 2011 to 2014 are presented as this period has gauged flows.

The Kling-Gupta Efficiency (KGE; Gupta et al., 2009) was used as an objective function for model calibration and performance evaluation to compare the simulated and benchmark flows. KGE was introduced to provide more reliable measures of the model performance (better estimation of flow variability) compared to the commonly used Nash-Sutcliffe Efficiency (Gupta et al., 2009). Each model (HYPE-ilake and HYPE-HDS) has separate model and calibration setups wherein the parameters in Table 1 were optimized within their permitted range to maximize the KGE objective function for the calibration period of SCRB only. Although the rating curve parameters are not applicable to the numerous small prairie depressions, we decided to include them in calibrating the HYPE-ilake model only to give the model a fair opportunity to replicate the streamflow to its best ability (ilratk and ilratp, Table 1). The Dynamically Dimensioned Search (DDS) algorithm (Tolson and Shoemaker, 2007) from the OSTRICH optimization toolkit (Shawn Matott, 2017) was used as the calibration algorithm for the study with a calibration budget of 10,000 optimization runs for each model.

2.5. Streamflow evaluation

The SCRB and SDNWA-90 simulated streamflows were compared to the observations using visual interpretations of the hydrographs and two qualitative metrics (KGE and PBIAS). KGE measures the goodness of fit for the overall hydrograph. Although KGE has a component that represents bias, we preferred to use PBIAS as a separate metric to measure the performance in preserving the overall runoff volume. It is believed that simulated flows having KGE > 0.4 and PBIAS < ±35% may be considered to be reasonable due to the difficulty in simulating this environment (Ahmed et al., 2021; Dibike et al., 2021). It might be seen that the ±35% is a significant value for runoff volume bias (Moriasi et al., 2007); however, it is typically considered an indicator of event capture problem rather than a mass balance issue (Ahmed et al., 2021).

3. Results

3.1. Streamflow performance

The streamflow hydrograph of both HYPE-ilake and HYPE-HDS as well as the gauged flows are shown in Fig. 3. For SCRB, HYPE-ilake showed a borderline satisfactory streamflow simulation during the calibration period but underestimated the total runoff volume (Table 2). Surprisingly, the performance of the model improved during the

Table 1

Calibration parameters, their description, range, and dependency for HYPE. HDS parameters were identified from GIS analysis of the topography.

| Name | Description | Units | min | max | Parameter dependency (number of instances) |
|----------|--|-----------|-------|-----|---|
| ttmp | Temperature threshold for snowmelt, snow density, and evapotranspiration | °C | −3 | 3 | Land cover (5) |
| cmlt | Melt factor | mm/°C/day | 1 | 10 | Land cover (5) |
| lp | Limit for potential evapotranspiration (PET) | – | 0.5 | 1 | General (1) |
| epotdist | Coefficient for PET depth dependency | 1/m | 2 | 6 | General (1) |
| cevp | Evapotranspiration parameter | mm/°C/day | 0.1 | 0.8 | Land cover (5) |
| bfrzon | Parameter for frozen soil infiltration | – | 1 | 5 | Soil type (2) |
| bcosby | Unfrozen soil coefficient | – | 1 | 20 | Soil type (2) |
| rrcs1 | Recession coefficient for soil layer 1 | 1/day | 0.05 | 0.6 | Soil type (2) |
| rrcs2 | Recession coefficient for soil layer 3 | 1/day | 0.001 | 0.2 | Soil type (2) |
| rrcs3 | Recession coefficient for slope dependence (soil layer 1) | 1/day/% | 0 | 0.1 | General (1) |
| wcwp | Wilting point fraction | – | 0.01 | 0.5 | Soil type (2) |
| wcfc | Field capacity fraction | – | 0.1 | 0.5 | Soil type (2) |
| wcep | Effective porosity fraction | – | 0.01 | 0.5 | Soil type (2) |
| sracs | Recession coefficient for surface runoff | 1/day | 0.01 | 0.2 | Land cover (4); set to 1 for depressions/lake |
| rivel | Water celerity | m/s | 0.1 | 10 | General (1) |
| damp | Delay coefficient that causes damping | – | 0.1 | 0.9 | General (1) |
| ilratk | Rating curve exponent for ilake outflow calculations | – | 1 | 150 | ilake region (1). Used by ilake module only. |
| ilratp | Rating curve coefficient for ilake outflow calculations | – | 1 | 5 | ilake region (1). Used by ilake module only. |

validation period in the hydrograph simulation (Fig. 3 and Table 2). Overall, the model was able to capture some peak flow events during the calibration period (e.g., 2001, 2011), but showed discrepancies in simulating the magnitude and/or timing of the rest of the events (e.g., 2010 and 2014) and completely missed some of the low flow events (e.g., 2007 to 2009 and 2015 to 2016, Fig. 3). The improvements of the model in the validation period likely occurred due to wetter (than the calibration period) conditions, which in this situation, the depressions would be full or near full, connecting the majority of the basin drainage network and maximizing the contributing area (i.e., no significant role of the variable contributing area, Ahmed et al., 2023).

Despite the reasonable simulation of the flows by the HYPE-ilake model at SCRB, the model was not able to simulate the SDNWA-90 hydrograph when transferring the parameters from SCRB as can be seen from Fig. 3 and the KGE score (Table 2). The model was not able to predict the accurate timing or magnitude of the 2011 event and it overestimated the spring snowmelt event of 2012 and the summer event of 2013 (Fig. 3). The model showed extreme discrepancies in replicating the runoff volumes at SDNWA-90 (PBIAS, Table 2). Clearly, the fixed lake threshold and fixed contributing area approach used to generate the runoff in this case is heavily reliant on calibration, meaning that the model can be successful at the calibration location but is more likely to

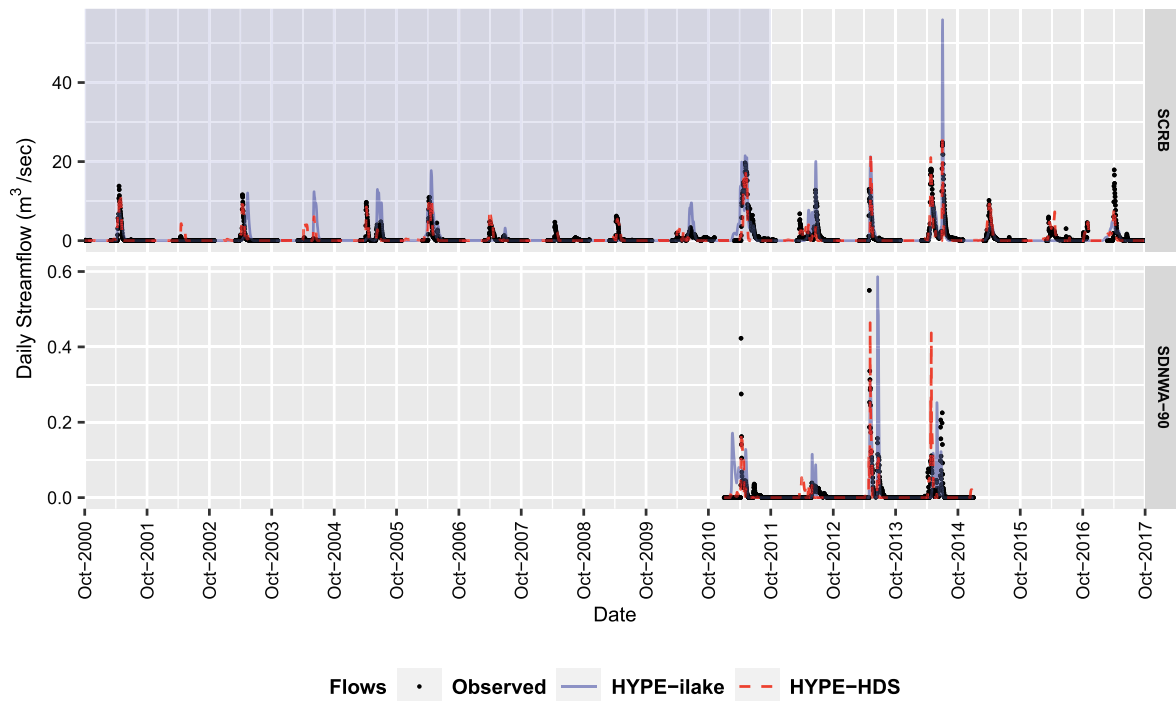


Fig. 3. Streamflow simulations of HYPE-ilake and HYPE-HDS for SCRB and SDNWA-90. The blue shaded area represents the calibration period, and the rest is the validation period. Each subplot has a different y-axis. SDNWA-90 was used as a validation site.

Table 2

Streamflow performance metrics of HYPE-ilake and HYPE-HDS for SCRB and SDNWA-90. Calibration metrics are not available for SDNWA-90 because this was used as a validation site.

| Basin | Model | KGE | | PBIAS (%) | |
|----------|------------|-------------|------------|-------------|------------|
| | | Calibration | Validation | Calibration | Validation |
| SCRB | HYPE-ilake | 0.42 | 0.61 | 39.84 | 0.50 |
| | HYPE-HDS | 0.83 | 0.67 | 2.21 | −21.10 |
| SDNWA-90 | HYPE-ilake | – | 0.02 | – | 52.92 |
| | HYPE-HDS | – | 0.59 | – | 0.15 |

fail when tested elsewhere because it does not have proper representation of the variable contributing area concept. This makes this approach inadequate for predicting the streamflow of the complex prairie region, especially for large-scale application when parameter transferability is typical.

The HYPE-HDS model showed improvement in the hydrograph simulation for SCRB compared to HYPE-ilake. HYPE-HDS showed good simulations of the hydrograph during the calibration period for SCRB (Fig. 3), which included the highest and lowest flow years. The model was able to replicate the overall hydrograph and was better able to represent high and low flow events, especially the 2011 flooding event. The model performance declined during the validation period showing reasonable performance during the 2014 flooding event, but overestimating the 2013 event (Fig. 3). HYPE-HDS simulated the hydrograph of SCRB quite well, based on the KGE scores for both the calibration and validation periods (Table 2). HYPE-HDS made better simulations of runoff volume during the calibration and validation periods, as shown by PBIAS values (Table 2), but still has large biases in the validation period. For SDNWA-90, HYPE-HDS showed reasonable spatial validation by using a transfer of parameters from another basin. The model was able to capture the 2011 and 2013 flooding event in that basin.

HYPE-HDS did not capture the summer event of 2014 and had problems in predicting the timing of the 2012 event (Fig. 3). The overall hydrograph simulation can be considered reasonable when looking at Fig. 3 and the performance metrics in Table 2. Biases can be due to a number of missing cold regions process representations in HYPE and are not completely modified by including the HDS algorithm.

HYPE-HDS outperformed HYPE-ilake model in terms of replicating the complex streamflow of two challenging prairie basins with very different landscapes (Fig. 3 and Table 2). However, HYPE-HDS was not able to replicate the accurate magnitude and/or timing of certain events in SDNWA-90, but still showed a better overall hydrograph simulation than HYPE-ilake. This is expected since the model was calibrated to fit the SCRB flows and the parameters were transferred to SDNWA-90. Furthermore, HYPE-HDS does not simulate gatekeeping, which occurs at SDNWA-90; therefore the errors in the SDNWA simulation are expected. Still, the results are very promising given that they were replicated for the non-conventional SDNWA-90 basin having no defined river network.

3.2. Storage and contributing area curves

The HYPE-ilake model is not designed to calculate the variable contributing area of the basin as it used a constant contributing area percentage of 100% throughout the simulation period. Therefore, it was necessarily excluded from the following contributing area analysis. The relationship between the fraction of stored water volume in the depressions (volume fraction) and the variable contributing area fraction of the basin was estimated using the HYPE-HDS model and is presented in this section.

3.2.1. SCRB

The relationship between the volume fraction and the variable contributing area fraction of the basin of HYPE-HDS for SCRB is plotted in Fig. 4 for each hydrological year. Fig. 4 shows a clear triangular (linear) hysteretic relationship between volume fraction and contributing area fraction. The curve consists of two phases: filling and

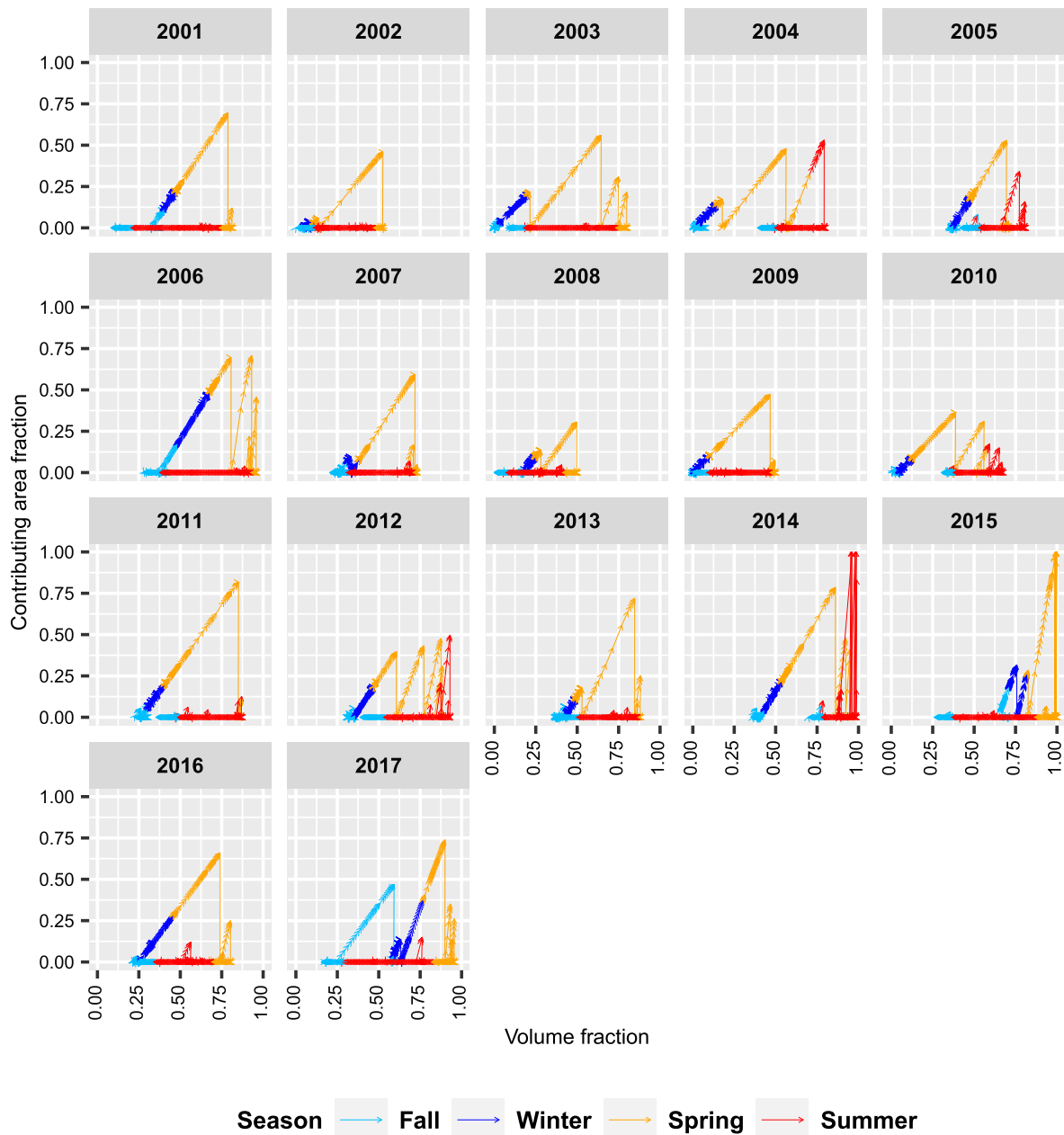


Fig. 4. The relationship between volume fraction and contributing area fraction of HYPE-HDS for SCR. Each subplot represents the relationship for a certain hydrologic year. Arrows show the direction of the relationship and colors represent the season.

emptying. Filling occurs when the depressions are being filled during snowmelt and rainfall events, due to net positive input to the depressions (Δd). This is depicted as the rising limb of the curve (hypotenuse of the triangles) where the depressions are being filled (increase in the volume fraction) and accordingly the contributing area increases. This can be seen in all hydrological years for SCR (Fig. 4). This rise occurred as more depressions are filled and begin to contribute flow to the outlet, unlike a traditional lake or water body that only surcharges its volume. These depressions are numerous (small lakes or water bodies), and they produce outflow when filled, but the relationship depicted in Fig. 4 is the result of the integrated depressions system when they work together as one system.

On the other hand, emptying of the depressions occur when water is being removed from the depressions due to evaporation and negative net input vertical fluxes. When this occurs, the contributing area has no role in the fill and spill process and it drops immediately to zero, regardless of

the volume fraction state (vertical drop in the figure, opposite sides of the triangles). This occurred because water was removed from all depressions and they were no longer filled or connected to each other. Accordingly, they stopped contributing to the main river, explaining why the contributing area dropped to zero. The contributing area remained at zero when the depressions were being emptied (flat horizontal line in all years, line parallel to the x-axis at zero contributing area). Note that this representation did not include large depressions and their gatekeeping role. A good graphical example of the state of filling and emptying depressions and the relationship with contributing area can be found in Clark and Shook (2022)'s study, Figs. 1 and 2 therein.

The contributing area depends on the history of additions and removals of water from the system. For small precipitation events, where the initial volume fraction and contributing fraction were small, the resulting volume fraction and the contributing area remained small (small leading triangles in 2002, 2003, 2004, and 2008; Fig. 4). It is also

possible for the volume fraction to be large, while the contributing fraction is small as in 2011 and 2017 (Fig. 4). In 2011, when the volume fraction was almost 0.9, (a small summer rainfall event, only increased the contributing area fraction to about 0.12, which shows that the history of changes in depressional storage, Δd , caused by net vertical fluxes and upland runoff had a greater effect on changing the contributing area. On the other hand, when significant amounts of water were added to the depressions (spring snowmelt in 2011; Fig. 4), the volume fraction increased significantly as the depressions filled and the contributing area increased substantially (almost 0.85) in a very short period of time (a couple of timesteps/days).

Fig. 4 also shows clear clockwise, nested, and consecutive loops (triangles) of hysteresis in the curves. This occurred when there were multiple filling (wet) and emptying (dry) events in the basin. However, dry years typically result in a smaller number of triangles (compared to wetter years) because the basin did not experience enough precipitation and/or runoff to cause multiple filling events; the only event producing runoff being the spring snowmelt (e.g., 2001, 2002, 2008, and 2009; Fig. 4). The year 2002 shows a single triangular loop, which occurred due to the spring snowmelt, followed by a relatively dry summer with low precipitation and little to no runoff leading to loss of water from depressions (small volume fraction) and little to no contributing area fraction (small leading triangle in 2002; Fig. 4). Wet years (e.g., validation period, 2012 to 2017; Fig. 4) showed more consecutive and nested loops compared to dry years (2001–2010) because the basin experienced multiple precipitation events (snowmelt and summer rainfall) that made the number of loops to increase drastically.

3.2.2. SDNWA-90

The shape and characteristics of the storage and contributing area relationship (explained earlier for SCRB) is very similar to that of SDNWA-90 (Fig. 5) in terms of direction, hysteresis, and linearity; however, they vary in their magnitudes. This is due in part to the differences in the meteorological forcings of the basins but is also related to the differences in their depressional storages. SDNWA-90 is dominated by large depressions, which have a greater ability to hold water. Therefore, the contributing area fraction never reached 1 for that basin, even when the depressional storage was greatest (e.g., 2011 and 2013; Fig. 5). Accordingly, the magnitudes of the streamflow generated by that basin are relatively small (Fig. 3).

3.3. Streamflow and contributing area timeseries for SCRB and SDNWA-90

The timeseries plots of the simulated streamflows and contributing

area fractions for SCRB and SDNWA-90 using HYPE-HDS are shown in Fig. 6. For SCRB, the contributing area timeseries are generally correlated with the streamflows (Fig. 6). It can be seen that the contributing area increased gradually during the snowmelt period (before the actual peak flow happens) as the depressions are filled by runoff from the melting snowpack. After this event, the contributing area returned to zero as water was removed from the depressions through evaporation. The contributing fraction remained very small during no flow (summer and fall) periods. However, when there were large rainfall events during summer periods, the contributing area increased rapidly (as the depressions had stored water from the snowmelt period, the soil is relatively wet, and the rainfall amounts were large) and flow can be generated (summers of 2004, 2012 and 2014, and fall of 2016; Fig. 6).

The magnitude of the streamflow is related to that of the contributing area. Flood years (2011 and 2014) had the greatest contributing areas, while low flow years had the smallest contributing area in SCRB (Fig. 6). However, for 2015, the contributing area remained high while the magnitude of flow was small (compared to the 2014 flood magnitude). This occurred because the basin was still wet from the large flood of 2014 and the depressions were nearly full. Therefore, when the snowmelt of the 2015 occurred (with relatively average snowpack depth), it caused the depressions to fill, making the contributing area of the basin (SCRB) nearly 100%. This is a demonstration of memory in the system, as described by Shook and Pomeroy (2011). In this case, the amount of net runoff reaching the river network was small even when the basin was fully contributing flow. The same can be seen for the SDNWA-90, peak flows correspond to relatively high contributing area (Fig. 6). However, in SDNWA-90, the contributing area never reaches 100%. This is true about this basin since it is dominated by relatively large depressions that have the ability to retain more water without being fully filled (Figs. 5 and 6).

4. Discussion

The HYPE-ilake model showed reasonable flow simulation during the calibration period, and its performance improved further during validation over the SCRB (Fig. 3 and Table 2). This was an artefact of the validation period being wetter than the calibration period, however, and in such situations, the fill and spill mechanism plays a more limited role in altering the runoff values as all depressions are nearly full and the majority of the basin can contribute flow to the outlet (Ahmed et al., 2023). The HYPE-ilake model failed when it was spatially validated at another test location (Fig. 3 and Table 2) because it lacks the physical representation of the variable contributing area complexities. This gives a good example of a model that is giving the right answer (reasonable

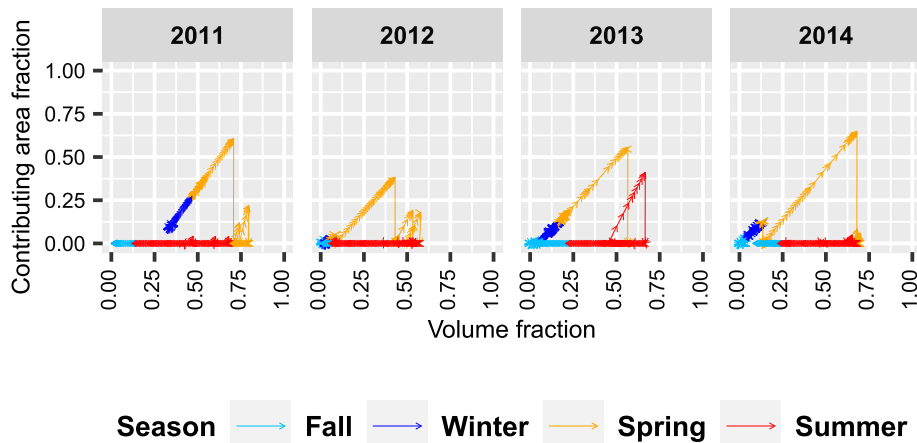


Fig. 5. The relationship between volume fraction and contributing area fraction of HYPE-HDS for SDNWA-90. Each subplot represents the relationship for a certain hydrologic year. Arrows show the direction of the relationship and colors represent the season.

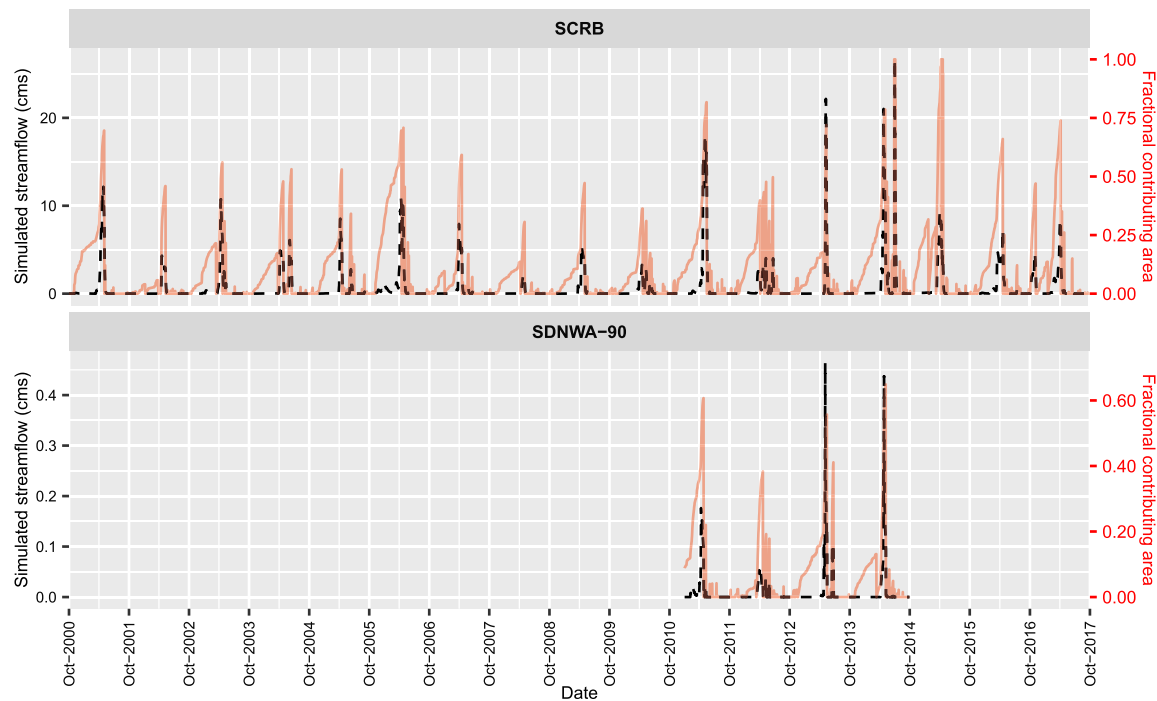


Fig. 6. Timeseries plot of the simulated streamflow and contributing area fraction generated by HYPE-HDS for SCR. Each subplot has different y-axis limits. The black dashed line represents the simulated streamflow timeseries (m^3/sec) while the solid light red line represents the fractional contributing area (—).

streamflow simulation) for the wrong reason (fixed contributing area) (Kirchner, 2006). This also suggests that the HYPE-ilake model heavily relies on calibration, and it might be successful at simulating select locations and time periods when rigorously calibrated, but its accuracy at other locations (or time periods) may deteriorate. This can limit the applicability of such a model in large-scale applications, and particularly simulation of ungauged basins or future time periods, wherein extensive calibration is not possible.

The HYPE-HDS model uses the concept of variable contributing area, which is a necessary requirement to simulate the prairie pothole fill-and-spill mechanism dominating the runoff response across the North American prairies. HYPE-HDS showed improved hydrograph simulation compared to HYPE-ilake. Also, HYPE-HDS maintained its good performance when spatially validated at another site, which occurred because the model could replicate the complex hysteretic relationship between contributing area and depressional storage, unlike HYPE-ilake (Figs. 4 and 5). A key similarity between HDS, ilake, and other traditional depressional storage algorithms that use probability distributions to represent the variable contributing area (e.g., HYPR, Ahmed et al., 2020b; SWAT, Mekonnen et al., 2016; and MESH, Mekonnen et al., 2014), is that they all use the concept of lumped (spatially integrated) depressional model, in which all depressions are represented as one storage unit. However, HDS has different underlying equations that allow it to replicate the known hysteretic relationships. Such hysteretic relationships cannot be replicated by traditional lake and depressional storage models (listed earlier) due to their limited formulation/parameterization, which is a key reason why they fail in predicting the complex prairie streamflow (Ahmed et al., 2021). Indeed, the failure of these models was the instigation for the development of the HDS algorithm.

The summer event of 2014 was one of the most widespread prairie floods that has occurred over the past decade. Many locations in the eastern Canadian prairies, such as SCR, were flooded and the entire basin was likely contributing flow at the outlet (Ahmed et al., 2021). HYPE-HDS simulation showed the same behavior as was observed in 2011 and 2014, where both the contributing area and volume storage

reached their capacity (Fig. 4). Also, during dry years (e.g., 2008 and 2010, SCR), the simulated contributing area was small compared to other years even though the depressions were more than half full. The agreement between the observations and HYPE-HDS suggest that the model can replicate aspects of the complex nature of the variable contributing area changes in the basin.

Overall, the hysteretic relationships depicted in Figs. 4 and 5 show how important it is to capture dominant lateral runoff generation processes occurring in prairie basins that lead to successful simulation of streamflows. The simulation curves are linear and simplified compared to the original curves produced by more complex and explicit models (Ahmed et al., 2020a, 2021; Shook et al., 2013; Shook and Pomeroy, 2011). However, this appears acceptable since the HDS conceptual approach was derived from more detailed explicit solutions (Shook et al., 2021). Further, this approximation/simplification allows integration of the HDS algorithm into almost any model, especially for large-scale domains when more complex approaches have limited applicability. The main intent of developing HDS is to have a depressional storage model with physically meaningful parameters that facilitates mapping parameters values to actual observations/measurements, which is particularly important for simulating large ungauged regions (e.g., pan-Arctic basin). Therefore, it is important to set the parameters values of HDS based on actual measurements or DEM data (Supplemental file, Section S1). This results in a more robust streamflow simulation as the model is constrained compared to adding additional calibration parameters (HDS parameters) that results in more degrees of freedom. Calibrating HDS parameters can result in significantly different process representation to achieve relatively similar streamflow performance as compared to setting HDS parameters values from observations (Supplemental file, Section S1). However, calibration can be a viable option if no reliable DEMs or observations are readily available, but that may be accomplished at the cost of model fidelity.

HYPE-HDS was proven to work with one of the most simplified forms of the HYPE modelling framework (one subbasin per watershed and conceptual calculations of the fluxes). This extreme test case shows the reliability of the HDS algorithm and its ability to work with more

simplified models in future. More discretization of the studied basins and using more complex and appropriate algorithms to calculate the vertical fluxes on the landscape and inclusion of the gatekeeping effect of large depressions (via activating lake module in HYPE), should improve HYPE-HDS ability in replicating prairie streamflow when they can be implemented.

5. HYPE-HDS software: model codes and implementations

The HYPE-HDS model/software was built based on the original HYPE model source code and uses the same input/output files and formats. The software was written in Fortran programming language and was successfully compiled using gfortran, tested, and run under Linux environment. The main modifications of the software were implemented in *model_hype.f90* and *sw_proc.f90* files from the original HYPE code to formulate the HYPE-HDS software by adding a few functions and sub-routines related to HDS functions/code. *modvar.f90*, *hypevar.f90*, *data.f90* and *model_hype.f90* files have been changed to allow reading the HDS algorithm new inputs from inputs files. *hypetypes.f90* and *assimilation_interface.f90* files have been changed to allow HYPE model to track changes in HDS state variables. Input data checks and tests were added to *hype_tests.f90* file.

To use the HDS algorithm within the HYPE modelling framework, the “*modeloption connectivity 2*” argument must be set in the *info.txt* file. Otherwise only original ilakes are simulated. The subbasins which should be simulated with HDS need to have an ilake SLC area fraction given in the subbasin information file (*GeoData.txt*). In this case, HYPE will activate HDS to run on that subbasin and treat the ilake SLC as a depressions SLC. The DS_{max} parameter of HDS can be specified per subbasin by adding a new column “*hds_depth*” in the *GeoData.txt* file or per ilregion (group/cluster of subbasins with the same depressional storage properties) by specifying the “*hdsdepth*” parameter in the parameter file (*par.txt*). To simulate subbasins without depressions, the user can set the *hdsdepth* parameter or ilake/depressions SLC fraction to zero. A detailed description of the model inputs files, units, and requirements are available on the comprehensive HYPE wiki website (SMHI, 2016). The development version of the source code is available on the following Zenodo repository (Ahmed et al., 2022). The long-term stable release of the code is available on the official HYPE website: <https://hypeweb.smhi.se/model-water/>.

6. Conclusion

The HYPE modelling framework was modified by adding a Hysteretic Depressional Storage (HDS) algorithm to represent the variable contributing area of a subbasin and to improve the streamflow simulation over the complex prairie region. HYPE-HDS was tested on two basins with very different landscapes namely, Smith Creek Research Basin (SCRB) and St Denis National Wildlife Area above pond 90 (SDNWA-90). Results showed that HYPE-HDS provided an improved simulation of the streamflows of both basins compared to current HYPE with its lake model. More importantly, HYPE-HDS can replicate the hysteretic relationship between the storage and contributing area, and the contributing area and streamflow. Such relationships cannot be predicted by the conventional HYPE-ilake model and can contribute to making the model unsuccessful when tested at other validation locations.

HYPE-HDS provides a more integrated solution to streamflow prediction problems in variably non-contributing regions, such as the prairie pothole region. More importantly, it can be tested for other modelling purpose such as sediment and nutrient transport, depressional ice and water temperature changes over prairie basins and Arctic permafrost depressions. This makes HYPE amongst a very few models that can simulate these important processes, which are crucial for improved agricultural practises in such regions. More importantly and given HYPE's ability to simulate reservoir and irrigation regulation, HYPE can now provide a step towards an integrated hydrological

modelling framework in the highly cultivated prairie area to simulate its streamflow, which will lead to better allocation and management of available water resources, more efficient hydropower generation, and greater water and food security in the Canadian prairies. Further efforts are needed to include the gatekeeping effects of large depressions and to allow for variable water area (depressional area) of the basin, which should further improve the simulation of the streamflow and evaporative fluxes at the basins scale. Further research is also needed to investigate the effect of the different DEM resolutions on the identification of the true depressions and to isolate and ignore false depressions.

Software and data availability

Software name: HYPE-HDS (HYdrological Predictions for the Environment - Hysteretic Depressional Storage).

Developers: Mohamed Ismaiel Ahmed [aut, cre], Kevin Shook [aut], Charlotta Pers [ctb, cre].

First year available: 2022.

Hardware requirements: PC/Mac.

Software requirements: Fortran compiler (e.g., gfortran).

Program language: Fortran.

Program size: 3.4 MB.

Software availability: <https://doi.org/10.5281/zenodo.7221439> (source code - development version); [hypecode.smhi.se](https://github.com/hypecode/smhi.se) (source code – long term stable version).

License: GNU LGPL-v3.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

A link to the software and data was shared in the paper.

Acknowledgment

The financial support of the Global Water Futures (GWF) project through GWF core modelling is greatly appreciated. The HYPE model was downloaded from the official HYPE website (hypecode.smhi.se). HYPE-HDS code (development version) and data used in this paper are publicly available at <https://doi.org/10.5281/zenodo.7221439>. HYPE-HDS code (recommended stable version) is available on the official HYPE website.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envsoft.2023.105769>.

References

- Agriculture and Agri-Food Canada, 2013. Areas of Non-contributing Drainage within Total Gross Drainage Areas of the AAFC Watersheds Project - 2013. <https://open.canada.ca/data/en/dataset/adb2e613-f193-42e2-987e-2cc9d90d2b7a>.
- Ahmed, M.I., Elshorbagy, A., Pietroniro, A., 2020a. A novel model for storage dynamics simulation and inundation mapping in the prairies. *Environ. Model. Software* 133, 104850. <https://doi.org/10.1016/j.envsoft.2020.104850>.
- Ahmed, M.I., Elshorbagy, A., Pietroniro, A., 2020b. Toward simple modeling practices in the complex Canadian prairie watersheds. *J. Hydrol. Eng.* 25, 04020024 [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001922](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001922).
- Ahmed, M.I., Elshorbagy, A., Pietroniro, A., Prinz, D., 2021. Improving the representation of the non-contributing area dynamics in land surface models for better simulation of prairie hydrology. *J. Hydrol.* 600, 126562 <https://doi.org/10.1016/j.jhydrol.2021.126562>.
- Ahmed, M.I., Shook, K., Pietroniro, A., Stadnyk, T., Pomeroy, J.W., Pers, C., Gustafsson, D., 2022. Implementing A Parsimonious Variable Contributing Area

- Algorithm for Prairie Pothole Region in the HYPE Modelling Framework. Zenodo. <https://doi.org/10.5281/ZENODO.7221439> [Software].
- Ahmed, M.I., Stadnyk, T., Pietroniro, A., Awoye, H., Bajracharya, A., Mai, J., Tolson, B. A., Shen, H., Craig, J.R., Gervais, M., Sagan, K., Wruth, S., Koenig, K., Lilhare, R., Déry, S.J., Pokorny, S., Venema, H., Muhammad, A., Taheri, M., 2023. Learning from hydrological models' challenges: a case study from the Nelson Basin Model Intercomparison Project. *J. Hydrol.*, 129820 <https://doi.org/10.1016/j.jhydrol.2023.129820>.
- Ameli, A.A., Creed, I.F., 2017. Quantifying hydrologic connectivity of wetlands to surface water systems. *Hydrol. Earth Syst. Sci.* 21, 1791–1808. <https://doi.org/10.5194/hess-21-1791-2017>.
- Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J.C.M., Hasan, A., Pineda, L., 2020. Global catchment modelling using World-Wide HYPE (WWH), open data, and stepwise parameter estimation. *Hydrol. Earth Syst. Sci.* 24, 535–559. <https://doi.org/10.5194/hess-24-535-2020>.
- Bajracharya, A.R., Ahmed, M.I., Stadnyk, T., Asadzadeh, M., 2023. Process based calibration of a continental-scale hydrological model using soil moisture and streamflow data. *J. Hydrol.: Reg. Stud.* 47, 101391. <https://doi.org/10.1016/j.ejrh.2023.101391>.
- Bajracharya, A., Awoye, H., Stadnyk, T., Asadzadeh, M., 2020. Time variant sensitivity analysis of hydrological model parameters in a cold region using flow signatures. *Water* 12, 961. <https://doi.org/10.3390/w12040961>.
- Bonsal, B.R., Aider, R., Gachon, P., Lapp, S., 2013. An assessment of Canadian prairie drought: past, present, and future. *Clim. Dynam.* 41, 501–516. <https://doi.org/10.1007/s00382-012-1422-0>.
- Clark, M.P., Shook, K.R., 2022. The numerical formulation of simple hysteretic models to simulate the large-scale hydrological impacts of prairie depressions. *Water Resour. Res.* 58, 1–25. <https://doi.org/10.1029/2022WR032694>.
- Costa, D., Shook, K., Spence, C., Elliott, J., Baulch, H., Wilson, H., Pomeroy, J.W., 2020. Predicting variable contributing areas, hydrological connectivity, and solute transport pathways for a Canadian prairie basin. *Water Resour. Res.* 56, e2020WR027984 <https://doi.org/10.1029/2020WR027984>.
- Dibike, Y., Muhammad, A., Shrestha, R.R., Spence, C., Bonsal, B., de Rham, L., Rowley, J., Evenson, G., Stadnyk, T., 2021. Application of dynamic contributing area for modelling the hydrologic response of the Assiniboine River Basin to a changing climate. *J. Great Lake. Res.* 47, 663–676. <https://doi.org/10.1016/j.jglr.2020.10.010>.
- Dumanski, S., Pomeroy, J.W., Westbrook, C.J., 2015. Hydrological regime changes in a Canadian Prairie basin. *Hydrol. Process.* 29, 3893–3904. <https://doi.org/10.1002/hyp.10567>.
- Evenson, G.R., Golden, H.E., Lane, C.R., D'Amico, E., 2016. An improved representation of geographically isolated wetlands in a watershed-scale hydrologic model. *Hydrol. Process.* 30, 4168–4184. <https://doi.org/10.1002/hyp.10930>.
- Fang, X., Pomeroy, J.W., 2009. Modelling blowing snow redistribution to prairie wetlands. In: *Hydrological Processes*, pp. 2557–2569. <https://doi.org/10.1002/hyp.7348>.
- Fang, X., Pomeroy, J.W., 2008. Drought impacts on Canadian prairie wetland snow hydrology. In: *Hydrological Processes*. John Wiley & Sons, Ltd., pp. 2858–2873. <https://doi.org/10.1002/hyp.7074>.
- Fang, X., Pomeroy, J.W., 2007. Snowmelt runoff sensitivity analysis to drought on the Canadian prairies. *Hydrol. Process.* 21, 2594–2609. <https://doi.org/10.1002/hyp.6796>.
- Fang, X., Pomeroy, J.W., Westbrook, C.J., Guo, X., Minke, A.G., Brown, T., 2010. Prediction of snowmelt derived streamflow in a wetland dominated prairie basin. *Hydrol. Earth Syst. Sci.* 14, 991–1006. <https://doi.org/10.5194/hess-14-991-2010>.
- Gasset, N., Fortin, V., Dimitrijevic, M., Carrera, M., Bilodeau, B., Muncester, R., Gaborit, É., Roy, G., Pentcheva, N., Bulat, M., Wang, X., Pavlovic, R., Lespina, F., Khedhaouria, D., Mai, J., 2021. A 10 km North American precipitation and land-surface reanalysis based on the GEM atmospheric model. *Hydrol. Earth Syst. Sci.* 25, 4917–4945. <https://doi.org/10.5194/hess-25-4917-2021>.
- Godwin, R.B., Martin, F.R.J., 1975. Calculation of gross and effective drainage areas for the Prairie Provinces. In: *Canadian Hydrology Symposium - 1975 Proceedings*, 11–14 August 1975, Winnipeg, Manitoba. Associate Committee on Hydrology. National Research Council of Canada, pp. 219–223.
- Gray, D.M., Landine, P.G., 1988. An energy-budget snowmelt model for the Canadian Prairies. *Can. J. Earth Sci.* 25, 1292–1303. <https://doi.org/10.1139/e88-124>.
- Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *J. Hydrol.* 377, 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>.
- Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C., Neal, J., 2022. A 30 m global map of elevation with forests and buildings removed. *Environ. Res. Lett.* 17, 024016 <https://doi.org/10.1088/1748-9326/ac4d4f>.
- Hayashi, M., Van Der Kamp, G., 2000. Simple equations to represent the volume-area-depth relations of shallow wetlands in small topographic depressions. *J. Hydrol.* 237, 74–85. [https://doi.org/10.1016/S0022-1694\(00\)00300-0](https://doi.org/10.1016/S0022-1694(00)00300-0).
- Hayashi, M., van der Kamp, G., Rosenberry, D.O., 2016. Hydrology of prairie wetlands: understanding the integrated surface-water and groundwater processes. *Wetlands* 36, 237–254. <https://doi.org/10.1007/s13157-016-0797-9>.
- Hayashi, M., Van Der Kamp, G., Rudolph, D.L., 1998. Water and solute transfer between a prairie wetland and adjacent uplands. 1. Water balance. *J. Hydrol.* 207, 42–55. [https://doi.org/10.1016/S0022-1694\(98\)00098-5](https://doi.org/10.1016/S0022-1694(98)00098-5).
- Hayashi, M., Van Der Kamp, G., Schmidt, R., 2003. Focused infiltration of snowmelt water in partially frozen soil under small depressions. *J. Hydrol.* 270, 214–229. [https://doi.org/10.1016/S0022-1694\(02\)00287-1](https://doi.org/10.1016/S0022-1694(02)00287-1).
- Kirchner, J.W., 2006. Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.* 42, 1–5. <https://doi.org/10.1029/2005WR004362>.
- Latifovic, R., Homer, C., Ressler, R., Pouliot, D., Hossain, S.N., Colditz, R.R., Olthoff, I., Giri, C., Victoria, A., 2016. 20 North American land-change monitoring system. *Remote Sens. L. Use L. Cover Princ. Appl.* 303.
- Lespina, F., Fortin, V., Roy, G., Rasmussen, P., Stadnyk, T., 2015. Performance evaluation of the Canadian precipitation analysis (CaPA). *J. Hydrometeorol.* 16, 2045–2064. <https://doi.org/10.1175/jhm-d-14-0191.1>.
- Lindsay, J.B., 2016. Whitebox GAT: a case study in geomorphometric analysis. *Comput. Geosci.* 95, 75–84. <https://doi.org/10.1016/j.cageo.2016.07.003>.
- Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S., 1997. Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.* 201, 272–288. [https://doi.org/10.1016/S0022-1694\(97\)00041-3](https://doi.org/10.1016/S0022-1694(97)00041-3).
- Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., Arheimer, B., 2010. Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Nord. Hydrol.* 41, 295–319. <https://doi.org/10.2166/nh.2010.007>.
- Mailhot, J., Bélair, S., Lefavre, 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., Qaddouri, A., 2006. The 15-km version of the Canadian regional forecast system. *Atmos.-Ocean* 44, 133–149. <https://doi.org/10.3137/ao.440202>.
- Martin, F.R.J., 2001. Addendum No. 8 to Hydrology Report #104, Agriculture and Agri-Food Canada PFRA Technical Service: Regina, Saskatchewan, p. 109. PFRA Hydrology Division, 1983, *The Determination of Gross and Effective Drainage areas in the Prairie Provinces*, Hydrology Repo.
- Mekonnen, B.A., Mazurek, K.A., Putz, G., 2016. Incorporating landscape depression heterogeneity into the Soil and Water Assessment Tool (SWAT) using a probability distribution. *Hydrol. Process.* 30, 2373–2389. <https://doi.org/10.1002/hyp.10800>.
- Mekonnen, M.A., Wheeler, H.S., Ireson, A.L., Spence, C., Davison, B., Pietroniro, A., 2014. Towards an improved land surface scheme for prairie landscapes. *J. Hydrol.* 511, 105–116. <https://doi.org/10.1016/j.jhydrol.2014.01.020>.
- Mengistu, S.G., Spence, C., Sprenger, M., Seeger, S., Blume, T., Weiler, M., 2016. Testing the ability of a semidistributed hydrological model to simulate contributing area. *Water Resour. Res.* 52, 4399–4415. <https://doi.org/10.1002/2016WR018760>.
- Moore, R.J., 2007. The PDM rainfall-runoff model. *Hydrol. Earth Syst. Sci.* 11, 483–499. <https://doi.org/10.5194/hess-11-483-2007>.
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., Binger, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* v. 50 <https://doi.org/10.13031/2013.23153>, 885-0-2007 v.50 no.3.
- Muhammad, A., Evenson, G.R., Stadnyk, T.A., Boluwade, A., Jha, S.K., Coulbaly, P., 2019. Impact of model structure on the accuracy of hydrological modeling of a Canadian Prairie watershed. *J. Hydrol. Reg. Stud.* 21, 40–56. <https://doi.org/10.1016/j.ejrh.2018.11.005>.
- Muhammad, A., Evenson, G.R., Stadnyk, T.A., Boluwade, A., Jha, S.K., Coulbaly, P., 2018. Assessing the importance of potholes in the Canadian Prairie Region under future climate change scenarios. *Water (Switzerland)* 10, 1–19. <https://doi.org/10.3390/w10111657>.
- Nasab, M.T., Singh, V., Chu, X., 2017. SWAT modeling for depression-dominated areas: How do depressions manipulate hydrologic modeling? *Water (Switzerland)* 9. <https://doi.org/10.3390/w9010058>.
- Phillips, R.W., Spence, C., Pomeroy, J.W., 2011. Connectivity and runoff dynamics in heterogeneous basins. *Hydrol. Process.* 25, 3061–3075. <https://doi.org/10.1002/hyp.8123>.
- Pomeroy, J., Fang, X., Westbrook, C., Minke, A., Guo, X., Brown, T., 2010. *Prairie Hydrological Model Study Final Report*.
- Pomeroy, J.W., Gray, D.M., Brown, T., Hedstrom, N.R., Quinton, W.L., Granger, R.J., Carey, S.K., 2007. The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence. *Hydrol. Process.* 21, 2650–2667. <https://doi.org/10.1002/hyp.6787>.
- Pomeroy, J.W., Gray, D.M., Landine, P.G., 1993. The prairie blowing snow model: characteristics, validation, operation. *J. Hydrol.* 144, 165–192. [https://doi.org/10.1016/0022-1694\(93\)90171-5](https://doi.org/10.1016/0022-1694(93)90171-5).
- Pomeroy, J.W., Shook, K., Fang, X., Dumanski, S., Westbrook, C., Brown, T., 2014. *Improving and Testing the Prairie Hydrological Model at Smith Creek Research Basin*.
- Shangquan, W., Dai, Y., Duan, Q., Liu, B., Yuan, H., 2014. A global soil data set for earth system modeling. *J. Adv. Model. Earth Syst.* 6, 249–263. <https://doi.org/10.1002/2013MS000293>.
- Shaw, D.A., Vanderkamp, G., Conly, F.M., Pietroniro, A., Martz, L., 2012. The fill-spill hydrology of prairie wetland complexes during drought and deluge. *Hydrol. Process.* 26, 3147–3156. <https://doi.org/10.1002/hyp.8390>.
- Shawn Matott, L., 2017. *OSTRICH – an Optimization Software Toolkit for Research Involving Computational Heuristics Documentation and User's Guide by L. Shawn Matott*, vol. 79. Ph. D. State University of New York at Buffalo Center for Computational Research.
- Shook, K., Papalexio, S., Pomeroy, J.W., 2021. Quantifying the effects of Prairie depressional storage complexes on drainage basin connectivity. *J. Hydrol.* 593, 125846 <https://doi.org/10.1016/j.jhydrol.2020.125846>.
- Shook, K., Pomeroy, J.W., Spence, C., Boychuk, L., 2013. Storage dynamics simulations in prairie wetland hydrology models: evaluation and parameterization. *Hydrol. Process.* 27, 1875–1889. <https://doi.org/10.1002/hyp.9867>.

- Shook, K.R., Pomeroy, J.W., 2011. Memory effects of depression storage in Northern Prairie hydrology. *Hydrol. Process.* 25, 3890–3898. <https://doi.org/10.1002/hyp.8381>.
- Spence, C., Woo, M.K., 2003. Hydrology of subarctic Canadian shield: soil-filled valleys. *J. Hydrol.* 279, 151–166. [https://doi.org/10.1016/S0022-1694\(03\)00175-6](https://doi.org/10.1016/S0022-1694(03)00175-6).
- SMHI, 2016. The HYPE wiki pages [HYPE Model Documentation] [WWW Document]. URL. <http://www.smhi.net/hype/wiki/doku.php>. accessed 2.21.23.
- Stadnyk, T.A., MacDonald, M.K., Tefs, A., Déry, S.J., Koenig, K., Gustafsson, D., Isberg, K., Arheimer, B., 2020. Hydrological modeling of freshwater discharge into Hudson Bay using HYPE. *Elem. Sci. Anthr.* 8 <https://doi.org/10.1525/elementa.439>.
- Tefs, A.A.G., Stadnyk, T.A., Koenig, K.A., Déry, S.J., MacDonald, M.K., Slota, P., Crawford, J., Hamilton, M., 2021. Simulating river regulation and reservoir performance in a continental-scale hydrologic model. *Environ. Model. Software* 141. <https://doi.org/10.1016/j.envsoft.2021.105025>.
- Tolson, B.A., Shoemaker, C.A., 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resour. Res.* 43, 1–16. <https://doi.org/10.1029/2005WR004723>.
- van der Kamp, G., Hayashi, M., 2009. Groundwater-wetland ecosystem interaction in the semiarid glaciated plains of North America. *Hydrogeol. J.* 17, 203–214. <https://doi.org/10.1007/s10040-008-0367-1>.
- Zeng, L., Shao, J., Chu, X., 2020. Improved hydrologic modeling for depression-dominated areas. *J. Hydrol.* 590, 125269 <https://doi.org/10.1016/j.jhydrol.2020.125269>.
- Zhang, B., Schwartz, F.W., Liu, G., 2009. Systematics in the size structure of prairie pothole lakes through drought and deluge. *Water Resour. Res.* 45 <https://doi.org/10.1029/2008WR006878>.