Modelling nutrient dynamics in cold agricultural catchments: A review

Diogo Costa\textsuperscript{a,b,c}, Helen Baulch\textsuperscript{a}, Jane Elliott\textsuperscript{c}, John Pomeroy\textsuperscript{a,b}, Howard Wheater\textsuperscript{a,b}

\textsuperscript{a} Centre for Hydrology, University of Saskatchewan, Canada
\textsuperscript{b} Global Institute for Water Security, University of Saskatchewan, Canada
\textsuperscript{c} Environment and Climate Change Canada, National Hydrology Research Centre, Canada

\textbf{A R T I C L E   I N F O}

\textbf{Keywords:}
Seasonal snow cover
Nutrient transport
Cold region processes
Hydrological controls
Catchment models

\textbf{A B S T R A C T}

The hydrology of cold regions has been studied for decades with substantial progress in process understanding and prediction. Simultaneously, work on nutrient yields from agricultural land in cold regions has shown much slower progress. Advancement of nutrient modelling is constrained by well-documented issues of spatial heterogeneity, climate dependency, data limitations and over-parameterization of models, as well as challenges specific to cold regions due to the complex (and often unknown) behaviour of hydro-biogeochemical processes at temperatures close to and below freezing where a phase change occurs.

This review is a critical discussion of these issues by taking a close look at the conceptual models and methods behind used catchment nutrient models. The impact of differences in model structure and the methods used for the prediction of hydrological processes, erosion and biogeochemical cycles are examined. The appropriateness of scale, scope, and complexity of models are discussed to propose future research directions.

\textbf{1. Introduction}

Agricultural activities and associated runoff of excess nutrients have impaired the ecological function of streams and lakes around the world (Carpenter et al., 1998; Withers and Lord, 2002; Schindler et al., 2012). This has led to significant efforts to reduce nutrient export and contain the growing global problem of eutrophication and cyanobacterial blooms. However, the quantification and prediction of nutrient exports to streams, lakes, and estuaries remain a difficult challenge despite decades of research on nutrient cycling and transport (Wade et al., 2008).

Cold climate regions are characterized by an average air temperature above 10 °C in the warmest months and below 3 °C in the coldest months (Peel et al., 2007). Cold regions hydrology is conceived as occurring in catchments where snowcover and frozen soils play a notable role in the hydrological cycle. Here, the problem of nutrient transport and pollution is affected by snow-related processes because they have a strong impact on flow generation, erosion and nutrient export (e.g. Brooks and Williams, 1999; Eimers et al., 2009; Casson et al., 2012). The spring freshet is often the major annual runoff event, and its magnitude and timing depend on both fall/winter processes and antecedent conditions, such as soil moisture, snowfall, and snow redistribution, as well as the characteristics of the snowmelt event such as duration, intensity and presence of frozen soils. In the Canadian Prairies, for instance, snowmelt runoff can account for more than 80% of the total annual runoff volume (Gray and Landine, 1988) and contribute the most nitrogen (N) and phosphorus (P) exported yearly (Corriveau et al., 2013). In these areas, snowmelt volume, melt rate and seasonally frozen soils are critical factors determining runoff-soil contact and erodibility (e.g. Ollesch et al., 2006; Panuska and Karthikeyan, 2010; Tiessen et al., 2010). Sub-zero temperatures, snowpacks, freeze-thaw cycling, and frozen soils may affect the biogeochemistry of these areas with impacts on nitrogen N and P.

There is considerable debate about the appropriateness of scale, scope, complexity, and accuracy of water quality models (Moore et al., 2006). Conventional process-based catchment nutrient models are increasingly complex and heavily parameterized but substantially simplify reality (Beck, 1987; Wade et al., 2008; Costa et al., 2019b). Uncertainties associated with hydrological and biogeochemical responses at various spatial scales, and sparse, sporadic water quality measurements with only rare measurements of key processes and pools further complicate the adequate use of catchment nutrient models, raising critical questions for the design, application, and benefit of such modelling tools.

This paper reviews the structure and conceptual foundation within widely used catchment nutrient models that have been applied in cold regions. It focuses primarily on processes specific to cold regions...
hydrology, which are the processes involving snow, ice and frozen soils as they affect the hydrological cycle. General hydrological and biogeochemical processes are also examined and discussed. Special attention is given to the processes directly affecting nutrient transport (magnitude, timing, and location), with an emphasis on the spring freshet as a period of great nutrient export to rivers and lakes. The review and model comparisons are used to (1) provide suggestions for model selection and recommendations for future research directions, and (2) discuss the appropriateness of scale, scope and complexity of nutrient models.

2. Nutrient export in cold agricultural regions: key processes

The movement of water and subsequent transport of nutrients in cold agricultural environments is strongly affected by the interplay of various snow, climate, soil and anthropogenic processes (Fig. 1). The relative importance of these processes varies depending on the location and land use.

2.1. Snowmelt and soil dynamics: hydrology, transport and erosion

Chemicals accumulate in snowpacks (winter) and soil, and snow processes, including snowmelt and infiltration into frozen soils, determine the hydrological pathways and residence times (Pomeroy et al., 2007) of these chemicals. Nutrient transport is often high during snowmelt (e.g., McConkey et al., 1997; Corriveau et al., 2013) due to frozen or thawing soils reducing infiltration and increasing runoff (Gaynor and Bissonnette, 1992; Gray et al., 2001). In the Canadian Prairies, for instance, most N and P is exported during this period (Corriveau et al., 2013). This is due to the duration and extent of snowmelt far exceeding that of the convective rainfall events that drive summer rainfall-runoff generation (Tiessen et al., 2010). Meltwater P and N may be infiltrated into the soil substrate, runoff overland, or follow a combination of surface and subsurface pathways, depending on the infiltrability of the soil (Granger et al., 1984), which may be affected by preferential infiltration of ions (Lilbaek, 2007), the presence of basal ice (Lilbaek and Pomeroy, 2008), concrete frost (Jones and Pomeroy, 2001), and macropores (Zhao and Gray, 1999), in addition to the soil physical characteristics and soil water content (Gray et al., 2001). Basal ice consists of an impermeable or patchy concrete frost that restricts and acts like a switch on infiltration (Gray et al., 1985; Jones and Pomeroy, 2001; Lilbaek, 2007). Its effect on flow generation and solute transport is complex and varies (Dunne and Black, 1971; Landon et al., 2004), but basal ice generally increases runoff efficiency by decreasing soil infiltrability and affects runoff chemistry by limiting the access to the soil matrix (Lilbaek, 2007).

A significant portion of nutrients can be stored and transformed in wetlands and topographic depressions (Neely and Baker, 1989; Johnston, 1991; Crompton and Isenhart, 1993; Birgand et al., 2007), and this storage may not necessarily connect to reach major streams and lakes in some regions (Spence et al., 2010; Shook et al., 1993). With a relatively

Fig. 1. Physical, chemical and biological processes affecting the accumulation and transport of nutrients in snow-covered areas.
impermeable layer of clay-rich glacial till deposits (Shaw and Hendry, 1998), wetlands in regions such as the prairie pothole region fill and dry in response to pluvial cycles and drought, and only spill in extremely wet conditions (Fang and Pomeroy, 2008). Snowmelt intensity and the presence of frozen soils determine the dominant infiltration mechanism (i.e., infiltration excess or saturation excess runoff) which has impacts on nutrient transport. Infiltration excess runoff occurs when what Gray et al. (1985) termed the “limited” case where soils are unsaturated, and water entry into the soil-ice-water matrix voids controls infiltration rates. Infiltration excess is often associated with large soil P losses (McDowell, 2012) because it often occurs when runoff begins, and the soil is still rich in nutrients. Additionally, thaw-freeze cycles reduce soil cohesive strength and increase soil erodibility, which promote sediment transport in early snowmelt (e.g., Edwards et al., 1995). This combined with frozen soils, which reduce soil permeability and increase runoff, often leads to large nutrient losses during this period. The relationship between the proportion of dissolved vs particulate nutrients and extent of soil frost is complex, varies during the melt period, and depends on the soil freeze-thaw dynamics. During early snowmelt, dissolved nutrients, both organic forms and inorganic forms (largely NO₃, NH₄, and PO₄), are transported in higher portions than particulates due to reduced erosion from frozen soils (Tiessen et al., 2010; Cade-Menun et al., 2013). Enhanced soil erodibility from freeze-thaw cycles often only occurs after the first spring thaw (Wall et al., 1988).

2.2. Snowmelt and soil dynamics: biogeochemistry

Cold climate conditions, as well as soil and nutrient management, have a strong impact on nutrient and sediment exports (Deelestra et al., 2009; Han et al., 2010). The nongrowing season is a critical period for nutrient losses in cold agricultural areas (Liu et al., 2019a), including those in Europe (Ulen et al., 2019), the Laurentian Great Lakes region (Good et al., 2019; Plach et al., 2019; Sadhukhan et al., 2019) and northern Great Plains of North America (Liu et al., 2019b).

N in soils is biologically influenced by processes such as mineralization and nitrification/denitrification, which are affected by oxygen (e.g., Bodelier et al., 1996), temperature and soil type and moisture (e.g., Saad and Conrad, 1993), as well as the availability and quality of organic matter (e.g., Breitenbeck and Brenner, 1987). Mineral NO₃ is a major N species which is readily transported with runoff and vulnerable to leaching (Snider et al., 2017) due to its high solubility in water and negative charge, so that it is not well retained by soil particles. The dynamics of P in soils (i.e., release, storage and speciation) depend on pH, temperature and organic carbon, and is affected by erosion and sorption to soil minerals and organic matter. The mass of P sorbed can increase with increasing concentrations of cations such as calcium Ca²⁺, magnesium Mg²⁺, and iron Fe³⁺ or Fe⁴⁺ in the soil (Haynes, 1984; Sanyal and De Datta, 1991; Khosravi et al., 2018), but frozen soils in cold regions can enhance the mobility of P by reducing runoff-soil contact and limit sorption rates in early spring snowmelt (Cade-Menun et al., 2013). Enhanced P release from prolonged runoff exposure to flooded acidic (e.g., Ann et al., 1999; Ajmone-Marsan et al., 2006; Scalenghe et al., 2012) and alkaline (e.g., Ponnamperuma, 1972; Amarawansa et al., 2015) soils is in sharp contrast to the otherwise often slow rate of soil nutrient release from soils (Dharmakeerthi et al., 2019; Schneider et al., 2019; Macrae et al., 2010). Amarawansa et al. (2015) examined 12 alkaline soils from Manitoba during summer conditions (laboratory incubation at 22 °C) and found, for instance, that P, Ca, Mg, and Mn were related. This suggested that P release was controlled by the dissolution of Mg and Ca phosphates and reductive dissolution of Mn phosphates. However, P release from flooded soils at low temperatures remains poorly understood despite its importance to nutrient transport during spring snowmelt (Amarawansa, 2013).

Low temperatures are generally associated with decreased rates of biological processing, including plant uptake, mineralization and nitrification/denitrification. However, high microbiological activity has been observed during snowmelt (Clark et al., 2009; Brooks et al., 1996; Tranter and Jones, 2001), in frozen and snow-covered soils (Peters and Driscoll, 1987; Brooks et al., 1996; Jones, 1999; Sebestyen et al., 2008; Clark et al., 2009; Pellerin et al., 2012; Snider et al., 2017), and in snowpacks (Jones and Deblois, 1987; Brooks et al., 1996; Mladenov et al., 2012) despite the low temperatures. Evidence of active biological processes has been found even in cold, barren and carbon-limited alpine soils (Williams et al., 1997, 2007; Brooks et al., 1999; King et al., 2008), arctic soils (Quinton and Pomeroy, 2006; Jones, 1999), as well as under lake ice (Cavaliere and Baulch, 2018). These winter and snowmelt processes can have important impacts on spring nutrient export. For example, mineralization of soil organic matter during winter has been linked to increased total dissolved P (TDP) availability during snowmelt (Freppaz et al., 2007) and reduced root uptake following freeze-thaw events has been associated with increased nutrient losses (Matzner and Borken, 2008).

The impact of freeze-thaw (FT) on nutrient release from soils (Liu et al., 2019a) and plant materials (Costa et al., 2019a; Liu et al., 2019a; Vanrobaeys et al., 2019), including crop residue and vegetated strips, is complex but has been receiving increasing attention (Liu et al., 2019a). Dormant (frozen) plants tend to cause higher nutrient losses during snowmelt runoff (Timmons et al., 1970), particularly P in actively growing young plants (Elliot, 2013). The maximum amount of biomass P released, which tends to occur in the first few hours of snowmelt (Costa et al., 2019a), depends on the extent of the plant cellular damage that is affected by the number of FT cycles, temperature of frost, and temperature-tolerance of the crop species (e.g., Bechmann et al., 2005; Ogaard, 2015; Cober et al., 2018; Costa et al., 2019a). Five FT cycles during a hard frost (temperature 18°C) have been observed to frequently cause the maximum nutrient release (Costa et al., 2019a; Bechmann et al., 2005).

Snow is also a reservoir of nutrients in snow-covered areas (Tranter and Jones, 2001) and snow N accumulated from atmospheric deposition can constitute a relevant source of readily available dissolved inorganic N (DIN, largely NO₃ and NH₄) during snowmelt (Jones, 1991). Although its contribution to the total N pool of soils is often small, soils contain primarily humic material not readily bio-available (i.e., organic N); thus inorganic N from snow can stimulate microbial and plant growth in early snowmelt (Jones, 1999). Ionic pulses in early snowmelt discharge are also common and may lead to temporary acidification of streams (Marsh and Pomeroy, 1999). This is caused by ice crystal metamorphism in snow forcing the reallocation of ions within the snowpack (Colbeck, 1976; Brimblecombe et al., 1985; Pomeroy et al., 2005; Lilbaek, 2007). It has been shown that this can cause the release of 50–80% of all snow ions within the initial 1/3 of the melt (Maule and Stein, 1990), with the process being potentially further exaggerated when meltwater runs over a basal ice layer (Hodson, 2006). Several laboratory and field studies have shown, however, that snow DIN can be quickly depleted in early snowmelt, particularly NH₄, via biological assimilation by snow algae (Jones and Soanchska, 1985; Jones and Deblois, 1987; Delmas et al., 1996).

Wetlands and agricultural ponds and reservoirs often retain runoff water under normal conditions (Fang and Pomeroy, 2008; Shook and Pomeroy, 2011), and they have the ability to attenuate nutrient exports (Price and Waddington, 2000; Fisher and Acreman, 2004; Tiessen et al., 2010) for a long time (e.g., Fernandes et al., 1996). However, they have a limited storage capacity (Helfield and Diamond, 1997) and certain factors may cause the release of soluble N and P species (Fisher and Acreman, 2004). In the Northern Great Plains, the processes controlling retention and release of nutrients are complex and poorly understood (Baulch et al., 2019a), requiring further research. Amongst the factors commonly related to the retention or release of nutrients in wetlands are (1) nutrient loading and duration, (2) hydraulic loading/retention time, (3) sediment oxygen/redox/water logging, (4) vegetation processes, (5) flow pathways, (6) fluctuating water table height, and (7) carbon...
content of wetland (Fisher and Acreman, 2004).

2.3. Agriculture practices

Agricultural practices affect nutrient transport if they alter hydrology, nutrient sources, or the interaction between water and nutrients during snowmelt runoff (Baulch et al., 2019b, for a detailed analysis). Nutrient transport is often higher during snowmelt, but nutrients applied to soils during the growing season may also be transported in rainfall-runoff events in the spring and summer (Nicholaichuk, 1967; Hansen et al., 2002; Glozier et al., 2006; Liu et al., 2013b). Nutrient management, primarily the addition of fertilizers or manure to support crop growth, is an important determinant of nutrient sources. The amount and form of nutrients applied, their placement (depth relative to runoff water penetration and location with respect to runoff pathways) and timing relative to runoff all contribute to the availability of nutrients for transport in runoff (Little et al., 2007). Tillage practices and perennial vegetation can affect runoff pathways and timing through their control on the height of stubble residue on the land during winter that impacts snow redistribution by blowing snow transport (Pomeroy and Gray, 1995; Pomeroy et al., 1993). Blowing snow redistribution to drainage channels and wetlands promotes snowmelt runoff amounts that exceed winter precipitation to these landscape units and makes them exceedingly important for runoff generation and in controlling streamflow generation processes. These practices can also influence infiltration and P stratification (Elliot and Efetha, 1999; Renton et al., 2015), as well as the amount of nutrients in soil and vegetation that can interact with runoff (Tiessen et al., 2010; Liu et al., 2014). Riparian buffer strips of perennial vegetation between cropland and streams will have a similar influence on blowing snow trapping and runoff generation. However, this influence is in a small but critical portion of the landscape that is very often the contributing area for runoff generation (Shook et al., 2015; Sheppard et al., 2006; Kieta et al., 2018). When perennial vegetation is grazed, nutrients are altered through consumption of forage and deposition of urine and feces (Haynes, 1984) and soil structure can deteriorate from compaction, limiting infiltration (Naeth et al., 1990). The pathway of runoff can also be affected by surface-/wetland or tile drainage, and this can affect the timing of runoff and exposure to nutrient sources (Brunet and Westbrook, 2012; King et al., 2015). For example, while surface flow quickly interacts with surficial soil layers and is able to transport nutrients located mainly in these regions, infiltration and tile flow interact with nutrients that may have leached through the soil profile through a more prolonged process.

3. Review of modelling methods

A relatively large number of water quality models are available. Mekonnen (2016) updated a previous compilation of water quality models by Shoemaker (1997) and identified 74 different models. The physical and biochemical principles underlining the methods used in each of these models, as well as the level of sophistication used in the methods deployed, frequently differ between models, often due to historical reasons (i.e., the initial motivation for developing the model). Also, catchment models require the representation of numerous processes, the importance of which varies between catchments. This often results in model developers representing processes differently depending on the model purpose and intended application.

Five models were selected for comparison in this study. The criteria used for model selection aimed at providing a wide-ranging overview of the different modelling philosophies and strategies. Selected models were: (1) catchment scale and (2) process-based. In addition, they support (3) long-term simulations and (4) simulate cold regions processes. For practical reasons, the analysis is limited to models that have been (5) widely used, (6) somehow tested in cold climates, and (7) for which adequate theoretical documentation is publicly available. Based on these criteria and taking into consideration a recent global overview and evaluation of watershed nutrient modelling presented by Wellen et al. (2015), a study which focused primarily on model performance comparisons and “best practices” in model implementation, the models selected for review in this study are as follows:

- **HYPE**: HYdrological Predictions for the Environment (Lindstrom et al., 2016; Arheimer et al., 2012)
- **HSFP**: Hydrological Simulation Program-Fortran (Bicknell et al., 2005; Duda et al., 2012), INCA and INCA-P: Integrated Model of Nitrogen and Phosphorus in Catchments (Whitehead et al., 1998; Wade et al., 2002, 2007; Jackson-Blake et al., 2016), SWAT: Soil and Water Assessment Tool (Arnold et al., 1998), and AnnAGNPS: The ANNUalized AGRicultural Non-Point Source (Bosch et al., 1998).

The analysis of the models is based on the latest official technical documentation available for each model: HYPE (March 2016), HSFP (September 1996), INCA (1998; INCA-P, July 2016), SWAT (September 2011), AnnAGNPS (March 2005). Refereed publications were also used to complement the information provided in the manuals. There is a version of SWAT developed for Canada (CanSWAT) that includes empirical algorithms for snow redistribution and frozen soils, as well as BMP modules for representation of small reservoir, holding pond, wetland conservation tillage, forage conversion, riparian grazing management (Yang, 2019). However, because information about the technical background and performance of CanSWAT is very limited, it could not be included in this review. The models selected for this analysis are compared regarding their (1) conceptual basis, (2) model structure, and (3) process representation. The objective of this study is to provide a general overview of the current modelling practices to highlight options for model application, further model development, and limitations to model implementation in some environments.

3.1. Conceptual basis

Catchment nutrient models have been primarily developed to support nutrient management. However, the purpose for (and the context in) which these models have been initially developed influences aspects related to model structure, theoretical foundations and temporal and spatial scales (Table 1).

Despite highlighting different modelling aspects in the official documentation, model applications show that they share the common purpose of supporting the evaluation of the effect of management decisions and climate change on water quality. HYPE, however, has been particularly designed with a focus on large, ungauged basins. All models are semi-distributed (e.g., utilize Hydrological Response Units, HRUs) and enable flexible temporal resolutions, but the daily timestep is a default setting for all models except HSFP that runs at hourly time intervals by default. Also, AnnAGNPS allows for fully-distributed (i.e., structured mesh) domain discretization. INCA and AnnAGNPS provide the simplest vertical discretization of the soil profile amongst the models examined, which consists of a 2-layer system where the upper layer corresponds to the reactive soil zone, and the lower layer corresponds to the deeper groundwater zone. The upper layer in AnnAGNPS is considered a tillage layer with a fixed thickness. At the opposite end of the spectrum is HSFP with the most complex soil structure - the soil is divided into four compartments with customizable properties. HYPE enables up to 3 layers and flexible parameterization.

3.2. Process representation

The computation of catchment processes in cold agricultural regions poses tremendous challenges due to hydrological, biogeochemical and management variability across seasons and landscapes. Winter and snowmelt processes impose specific challenges to measuring and
characterizing processes such as snow sublimation, blowing snow redistribution, snow water content, snowmelt runoff, frozen soil, freeze-thaw cycles and overwinter biogeochemical cycling. In this section, the methods implemented in the different models for calculation of processes of (1) general hydrology, (2) cold regions hydrology, and (3) biogeochemistry are reviewed and compared. Focus is given to aspects of hydrology and biogeochemistry during the winter and snowmelt period, with the broad topic of general hydrology (i.e., rainfall-runoff) being only briefly covered. The results are summarized in easy-to-read circular plots, with the full review material used to produce these

Table 1
Purpose, model structure and default temporal resolution of the models. The “Purpose as highlighted in official documentation” was taken verbatim from the respective technical manuals.

<table>
<thead>
<tr>
<th>Purpose (or original purpose) as highlighted in the official documentation (this information was taken verbatim from the manuals)</th>
<th>HYPE</th>
<th>HSPF</th>
<th>INCA</th>
<th>SWAT</th>
<th>AnnAGNPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose as highlighted in official documentation (this information was taken verbatim from the manuals)</td>
<td>Prediction of Ungauged Basins</td>
<td>Assess the effects of land-use change, reservoir operations, point or nonpoint source treatment alternatives, flow diversions, etc.</td>
<td>Prediction of environmental change issues in catchments including land use change, climate change and changing pollution environments including point and diffuse pollution</td>
<td>Prediction of the impact of land management practices in large, complex watersheds</td>
<td>Evaluating the effect of management decisions impacting water, sediment and chemical loadings within a watershed.</td>
</tr>
<tr>
<td>Representation of space Vertical elements (details on these properties are provided in sub-Section 3.2)</td>
<td>Integration of water and water quality. Large model setups</td>
<td>All models are semi-distributed, but AnnAGNPS may also be used as a fully-distributed model</td>
<td># Layers: Up to three soil layers</td>
<td># Layers: Two soil layers</td>
<td># Layers: Up to 10 soil layers</td>
</tr>
<tr>
<td>Properties: properties can be fully and independently defined for each layer</td>
<td>Properties: properties can be fully and independently defined for each layer</td>
<td>Properties: simple two box approach - a reservoir for the reactive soil zone and another reservoir for the ground-water zone.</td>
<td>Properties: Ten-box soil model with two groundwater aquifers (shallow and deep) underneath based on a linear reservoir approximation</td>
<td>Properties: The top layer is considered a tillage layer with constant thickness (20 cm) unless bedrock is within 20 cm of the surface. The second layer has static properties and does not exceed a 2-m depth from the surface.</td>
<td></td>
</tr>
<tr>
<td>Default time step</td>
<td>Daily</td>
<td>Hourly</td>
<td>Daily</td>
<td>Daily</td>
<td>Daily</td>
</tr>
</tbody>
</table>

Fig. 2. Process representation in the models for general (left a) and cold regions’ (panel b) hydrology. The rings represent different models from HYPE in the outer ring to AnnAGNPS on the inner circle. Processes are described for each slice outside the circle, and where that process is represented, the box is coloured. White boxes mean that the process is not represented in the model. Codes inside the small circles (e.g., A1, D7) are respectively the row and column in Table A1 in Supplementary Material from where the information was taken to produce the displayed results.
graphics provided as Supplementary Material, to facilitate model inter-comparisons.

3.2.1. Hydrological processes

Adequate (background) hydrological-hydraulic modelling is a critical step for successful nutrient transport simulations. Fig. 2 summarizes the modelling capabilities of the reviewed models regarding both general (panel a) and cold regions (panel b) hydrology. Due to the diversity of the methods deployed in these models, the methods are classified as (1) physically-based or with more sophisticated process representation, (2) semi-empirical or with intermediate sophistication in process representation, or (3) empirical. The reader is referred to Supplementary Material for access to all the detailed information compiled about each model that was used to produce the figure.

AnnAGNPS has overall the most comprehensive and physically-based representation of both general and cold regions processes, and INCA has the simplest. Evapotranspiration is the only general hydrology process (panel a) that is simulated with similar type of methods across all models (i.e., Penman-Monteith and other comparable methods). Groundwater and erosion are either neglected or simulated based on empirical methods in all models, although differences exist in the level of detail and complexity of the approaches used. While it is common that groundwater is simplified in mostly hydrological models, the sensitivity of groundwater representation can vary markedly across regions (Panagoulia and Dimou, 1996; Erickson and Stefan, 2009; Carey et al., 2013). Although erosional transport of particulate P is important in most landscapes (Tiessen et al., 2010; Su et al., 2010), modelling erosion remains a major scientific challenge (Fu et al., 2019) and most models rely on parametric estimations of kinetic energy from raindrops and surface runoff to predict erosion, with SWAT and AnnAGNPS applying the popular RUSLE method (Renard et al., 1991) for erosion. Tile drainage drainage is simulated in most of the models based on the popular empirical Hooghoudt equation or other simplified parametric expressions.

AnnAGNPS is the most sophisticated and physically-based model concerning cold regions processes. The model can simulate most of the processes examined primarily because it solves the full energy balance across the snowpack and soil layers, which is used to estimate snowmelt, runoff over frozen soils, rain-on-snow and erosion of soil layers. These processes combined have been recognized as an important control of nutrient export at field scales (Costa et al., 2017; Costa and Pomeroy, 2019), and the importance of their physically-based representation has long been documented (Pomeroy et al., 2007). One process that is neglected in AnnAGNPS is the wind redistribution of snow, a process that can significantly affect the spatial patterns of snowmelt runoff generation in open windswept agricultural regions, such as the northern US Great Plains, and the Canadian Prairies in North America (e.g., Pomeroy and Gray, 1995). While none of the models simulate blowing snow explicitly, HYEP and SWAT have simple approaches to estimate snowcover heterogeneity (i.e., sub-grid variability of late-season snow-water-equivalent (SWE)) based on elevation and land use (in the case of HYEP) and areal depletion curves based on elevation bands (in the case of SWAT). INCA, in turn, is on the other side of the spectrum as to the simulation of cold regions processes. It provides the most straightforward model framework largely based on parametric or empirical methods.

Snowfall, snow accumulation, snowmelt and soil temperature are calculated in all models, but the methods used vary. Snowmelt is calculated using the empirical temperature-index method in most models, except for HSFP and AnnAGNPS, which compute the snowpack energy balance. The use of the temperature-index method can be problematic because it requires recalibration for every new regional climatic input (Walter et al., 2004), requires different parameters for rain-on-snowmelt, and does not work well where snowmelt is dominated by solar radiation inputs (Male and Gray, 1981). The soil temperature is computed based on empirical relationships with air temperature in most models. AnnAGNPS is the exception as it obtains this information from the energy balance computations.

The reader is referred to Supplementary Material for a more detailed description of the methods used in each model.

3.2.2. Biogeochemistry (nutrient pools, sources/sinks, and biogeochemistry)

3.2.2.1. Soil nutrient pools. Table 2 summarizes the nutrient pools used to model the different N and P species simulated in each model. The conceptual basis used to represent N and P speciation varies significantly across models, with HSFP and SWAT emerging as the models with the most complex partitioning of mineral and organic N and P in soils (8 and 6 pools for N and P, respectively) and INCA with the simplest (2 and 3 pools for N and P, respectively).

In the case of mineral N, HYPE and AnnAGNPS lump all species as dissolved inorganic nitrogen (DIN), despite this obscuring the different specific controls on the fate of NH4 and NO3. The remaining models differentiate between NO3 and NH4, with HSFP further subdividing NH4 into soluble and particulate-bound fractions. In the case of organic-N, DON is only simulated by HYPE and HSFP, with the latter subdividing it further into labile and refractory fractions. Although DON is an overlooked pathway of nitrogen loss, it can be the dominant nitrogen species exported in agricultural systems, with the environmental effects of urea, for instance, raising growing concerns (e.g., Donald et al., 2013, 2011). HYPE provides the simplest conceptual model for simulation of organic-N that is based on DON, residue, and labile-PON (named as “fastN”) and refractory-PON (named as “slowN”) pools. AnnAGNPS, SWAT and HSFP divide particulate organic-N (PON) into three pools, i.e., fresh residue and labile and refractory PON. INCA does not simulate organic-N explicitly, including its effect through sink and source terms. This is a limitation since organic N can be reactive, even in cold climates (e.g., Chantigny et al., 2019), and also accumulate in agricultural areas through the continuous use of excess fertilizer, which can cause large accumulation of organic N in soils and groundwater that can persist for decades (Van Meter et al., 2018).

Substantial differences also exist in the way P is simulated in the reviewed models. INCA-P has the simplest conceptual model comprising of 3 pools, a TDP pool, a mineral labile-P pool and a combined/lumped inactive organic and mineral P pool. HYPE and HSFP are more detailed and divide mineral-P into SRP and particulate-bound P, while AnnAGNPS and SWAT go further and include an additional pool for mineral non-active soluble-P. The organic P pools in HYPE are DOP, “fastP” (which appears to comprise both residue-P and labile-POP) and “slowP” (representing refractory-POP). DOP is only modelled in HYPE, but all the remaining particulate fractions of organic P are simulated by the remaining models. However, there are important differences in the way the particulate organic pools (POP) are sub-divided. AnnAGNPS and HSFP have similar approaches, which consist of differentiating between fresh residue-P and humic-P. SWAT uses a similar method but further divides humic-P into active and stable sub-pools. Finally, HYPE uses a slightly different conceptual model where plant residue and the active portions of humic-P are lumped into one single pool (organic “fastP”), with the remaining organic-P being classified as “slowP”, which stands for “slow (reacting organic) P.”

The approaches used in the models to group the different nutrient species into shared pools has implications on the type and number of transformation pathways that are explicitly represented. To a large extent, this determines the scope of the model since it defines the biogeochemical processes that can be included in the simulations. However, increasing the number of pools and overall complexity of biogeochemical pathways imposes significant challenges. More nutrient pools and biogeochemical pathways require additional model parameterization, a procedure that is often problematic because of little supporting field data. This suggests that simpler biogeochemical modelling
Table 2
Nutrient pools represented in the models: N (left panel) and P (right panel).

(a): N cycle

<table>
<thead>
<tr>
<th>Mineral</th>
<th>Organic</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIN: dissolved inorganic N</td>
<td>DON: dissolved organic N</td>
</tr>
<tr>
<td>NO₃: nitrate</td>
<td>DONlab: labile DON</td>
</tr>
<tr>
<td>NH₄⁺: ammonium</td>
<td>OrgNDONlab: refractory DON</td>
</tr>
<tr>
<td>NH₄⁺ solution: NH₄ in solution</td>
<td>OrgN DONrefractory: labile DON</td>
</tr>
<tr>
<td>NH₄ part-absor: particulate-bound NH₄</td>
<td>OrgN DONlab: particulate-bound organic N</td>
</tr>
</tbody>
</table>

(b): P cycle

| TDP: total dissolved P |
| SRP (MinerPolution-refractory): soluble reactive P |
| PO₄-labile: labile phosphate |
| PO₄-absor: particulate-bound PO₄ |
| partP (MinerPpart): particulate-bound mineral |
| partP-labile (MinerPpart-labile): labile particulate-bound mineral P |

Table 3
Biogeochemical transformations of soil N calculated by each model. The transformations are grouped into 7 reaction types: denitrification (A), nitrification (B), sorption/immobilization (C), desorption/dissolution (D), mineralization (E), organic decomposition (F) and synthesis/return (G).

<table>
<thead>
<tr>
<th>MINERALS</th>
<th>ORGANIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIN: dissolved inorganic N</td>
<td>DON: dissolved organic N</td>
</tr>
<tr>
<td>NO₃: nitrate</td>
<td>DONlab: labile DON</td>
</tr>
<tr>
<td>NH₄⁺: ammonium</td>
<td>DONref: refractory DON</td>
</tr>
<tr>
<td>NH₄⁺solution: NH₄ in solution</td>
<td>PON: particulate organic N</td>
</tr>
<tr>
<td>NH₄ part-absor: particulate-bound NH₄</td>
<td>PONlab: labile PON</td>
</tr>
<tr>
<td>N2: molecular N</td>
<td>PONref (humusN in HYPE): refractory PON</td>
</tr>
<tr>
<td>NH₄part-absor: particulate-bound NH₄</td>
<td>ONres: ON from plant residue</td>
</tr>
</tbody>
</table>
calibration instead of on observable, transferable and regionalized parameters.

3.2.2.2. Soil N biogeochemistry. Table 3 shows the biogeochemical transformations of soil N represented in each model, which are limited by the nutrient pools (i.e., model state-variables) they simulate. For the sake of model inter-comparison, the transformations pathways are divided into 7 groups based on reaction types: (1) denitrification, (2) nitrification, (3) sorption/mobilization, (4) desorption/dissolution, (5) mineralization, (6) organic decomposition (within organic species), and (7) organic synthesis (from both mineral and organic species). Organic decomposition is used to group all the transformations that involve going from more stable (more complex) species to labile (simpler) organic N forms, with the opposite being designated as synthesis/return. Dissolution is used to describe the processing where PON dissolves in water transforming into DON.

HSPF has the most complex biogeochemical model structure, enabling a total of 11 reaction pathways (from 7 transformation groups) between its 8 N pools. In turn, INCA and HYPE have the simplest model structure providing five possible reaction pathways (from 4 transformations groups). HYPE and AnnAGNPS simulate DIN without differentiating between NO\(_3^-\) and NH\(_4^+\) (rows A and B in the table). Hence nitrification (transformation of NH\(_4^+\) into NO\(_3^-\)) is not simulated, and denitrification is indirectly computed from a pre-determined (i.e., hard-coded) fraction of DIN that is intended to represent the proportion of NO\(_3^-\) in DIN. This obscures the different controls on the fate of NH\(_4^+\) (e.g., volatilization) versus NO\(_3^-\) (e.g., denitrification). The remaining models simulate NO\(_3^-\) and NH\(_4^+\) explicitly. Mineralization is computed by all models, although the number of transformation pathways for this type of reaction varies (row E). HYPE and HSPF enable only one mineralization pathway each: from fastN (i.e. labile-N) to DIN, in the case of HYPE, and from PON-labile to NH\(_4^+\) in the case of HSPF. INCA accounts for mineralization through source terms, while AnnAGNPS and SWAT further accounting for the transformation of residue-ON into refractory-PON. On the other hand, SWAT and HSPF enable the simulation of residue-ON and refractory-PON into labile-PON, with AnnAGNPS further accounting for the transformation of residue-ON into refractory-PON. Finally, synthesis/return within organic N species (row G, the inverse of decomposition in row F) is only simulated within AnnAGNPS and HSPF through labile-PON into refractory-PON transformations.

Fig. 3 shows the controls of the different N reactions included in the models reviewed. All transformations are represented as first-order reactions that depend on the N species consumed (e.g., nitrate in the case of denitrification) and a decay parameter estimated or calibrated based on different environmental controls. The impact of temperature on reaction rates is included in all models and for all transformations. The impact of soil moisture is also taken into account in most models and most reactions, except HSPF, which only enables limiting nitrification for soil moisture. However, the effect of substrate limitation on denitrification is only considered by SWAT and HYPE, which use organic carbon concentrations and half-saturation constants (or the Michaelis constant) for application of the Michaelis–Menten kinetics, respectively. Half-saturation constants are also used to limit sorption rates in HYPE and HSPF.

3.2.2.3. Soil P biogeochemistry. Table 4 compares the biogeochemical transformations for soil P represented in each model. Similarly to the N models, the transformations of P species are grouped into reaction types: (A) sorption/desorption, (B) mineralization, (C) decomposition and weathering, (D) immobilization (mineral to organic), (E) synthesis/return between mineral species, (F) dissolution, (G) decomposition, and (F) synthesis/return between organic species. Also here, organic and mineral decomposition is used to group all the transformations that
involve going from more stable (more complex) organic N species to labile (simpler) forms, with the opposite being designated as synthesis/return. The term “Dissolution” is used to describe the process where POP dissolves in water transforming into DOP; the term “Mineralization” includes all transformations that convert an organic P species into a mineral P species; the term “Decomposition” within mineral or organic species is used to group all the transformations that alter a stable/refractory P species into a reactive/labile one; and the term “weathering” is only used to explicitly include this designation as used in INCA to characterize processing of InactiveP into LabileP. INCA-P and HSPF have the simplest biogeochemical model for P; SWAT and AnnAGNPS have the most elaborate. Sorption/desorption dynamics are simulated by all models (row A), but there are significant differences in the methods deployed. While HYPE and HSPF use the Freundlich isotherms, SWAT uses a P availability index, INCA uses an updated version of the equivalent P concentration which employs a dynamic variable calculated as a function of adsorbed P, and AnnAGNPS determines this equivalent dynamics as a function of the soil sediment clay fraction and an empirical partitioning coefficient. Mineralization is computed by all models (row B), with HYPE and HSPF providing the simplest pathway from POP to SRP, and SWAT allowing for additional mineralization routes from the different forms of POP (POPres and POPact) into the various species of mineral-P: SRP, MinerPsol, partP and/or PO4act depending on the model and the pools it simulates. In turn, mineral-to-organic immobilization (row D) is only simulated by HSPF via active/ soluble PO4 (POPact) conversion into POP.

Decomposition (row C) and synthesis/return (row E) within mineral-P species are only simulated by SWAT and AnnAGNPS through a series of transformation pathways from/to (decomposition/synthesis) partP and MinerPsol to/from (synthesis/decomposition) SRP. HSPF only allows the processing of PO4act into POP (immobilization, row D), and INCA takes a more generic approach based on exchanges between labile-P and inactive-P pools.

The cycling of P within organic forms (rows F-J) is only simulated by HYPE, AnnAGNPS, and SWAT, with INCA treating all organic-P as a single pool named inactive-P (rows D and E at the INCA column). While AnnAGNPS has the simplest conceptual model for P, one which is limited to the decomposition of residue-P into POP (row G), SWAT has the most complex conceptual model for P, which includes the decomposition of POPres into POPact and POPstab, and of POPstab into POPact (row G), as well as synthesis/return of POPact into POPstab (row J). HYPE only allows for the decomposition of “fastP” (corresponding to POPres and POPact together) into “slowP” (corresponding to POPstab). The processing of POP (corresponding to “fastN” and “slowN” in HYPE) into DOP is only considered in the HYPE model (row F).

Fig. 4 shows the controls of the reaction rates included in the models reviewed. Similar to the N cycle, the temperature dependency is the environmental factor most commonly used across the different models to control the rate of the various transformation pathways within the P cycle, except for sorption/desorption (slice A). Here, sediment properties are used in all models for application of the Freundlich isotherm or other methods (see Table 4). Soil moisture is also used in HYPE, HSPF, and AnnAGNPS to control mineralization rates (slice B), as well as decomposition rates (slices C and G) in HYPE and SWAT, sorption/desorption (slice A) in HSPF, and processing of POP into DOP (slice F) in HYPE. AnnAGNPS uses a rather complex approach to simulate decomposition within mineral species (partP into SRP and MinerPsol into SRP, slice C) through a piece-wise function that depends on soil pH levels. Here, factors such as temperature, soil moisture, the concentration of CaCO3 and organic carbon are selectively combined to compute
decomposition rates depending on pH levels.

The reader is referred to Supplementary Material for more information about the data compiled to generate Tables 2–4. There, details are also provided regarding the types of nutrient sources enabled in each model.

4. Discussion

4.1. Model limitations and strengths: suggestions for model selection and recommendations for future directions

The models reviewed in this study exhibit conceptual differences that can be important depending on the region and application. These differences are related to the characterization of the case study domain (horizontal and vertical computational elements) and the methods deployed to simulate the different hydrological and biogeochemical processes. In this section, various modelling aspects are discussed as to their strengths and limitations for different model applications. This is used to provide suggestions for model selection and advance recommendations for future research.

4.1.1. Model structure: the problem with heavily stratified soils

The models reviewed are all semi-distributed (e.g., Hydrological Response Units, HRUs) and generally divide the soil into between two to four layers, except for SWAT that allows for up to 10 layers, and AnnAGNPS that can also be used as a fully-distributed model (i.e., structured mesh). While the use of few vertical layers is common practice in hydrology (e.g., HYPE, TOPKAPI (Ciarpica and Todini, 2002)), additional surficial soil layers may be needed in heavily stratified agricultural soils (i.e., subject to tillage) to adequately represent the accumulation of fertilizer or manure application and soil mixing practices, which determine the opportunity of soil nutrients to interact with runoff. It has been shown that the amount and form of nutrients applied, their placement (depth relative to runoff water penetration) and timing relative to runoff all contribute to the availability of nutrients for transport in runoff (Little et al., 2007). Tillage practices can influence soil stratification, runoff pathways and infiltration (Elliott and Efetha, 1999; Renton et al., 2015) and the amount of nutrients in soil and vegetation that can interact with runoff (Tiessen et al., 2010; Liu et al., 2014). Coarse vertical soil resolutions in models also require the averaging of detailed soil data from soil surveys for use as model input/forcing or for validation, which may lead to the loss of important information to understand the temporal and spatial scale of processes.

4.1.2. General hydrology: call for meaningful model structures based on observable and transferable parameters

There are several differences in the representation of hydrological processes between the models, differences which may have important implications for nutrient transport predictions. The modelling of rainfall-runoff-infiltration is generally based on simplified empirical methods such as the SCS Curve number approach, techniques based on saturation and infiltration excess values or residence times and baseflow index concepts, with only SWAT using the Green & Ampt method for infiltration and the rational method for peak runoff rates. Similarly, the routing schemes for streamflow used are primarily based on basic hydraulic principles that range between empirical and power-law based functions, depth-area-volume-flow relationship tables and variable storage, and the Muskingum method with the Manning’s Equation for open channel flow. While some level of empiricisms and parameterization is often needed for more efficient computations at catchment-scales, meaningful model structures that combine methods based on observable and transferable parameters are needed. We argue that methods like the SCS Curve number approach for rainfall-runoff estimation and empirical or power-law formulations for routing, for instance, are undesirable for they rely on non-observable, less-transferable and non-regionalized model parameters.

4.1.3. Cold regions’ hydrology: most models decades behind the science

The hydrology of cold regions is strongly influenced by the seasonality of air temperature and soil and snow energy balances, snow
precede unfrozen ones, and particulates are often depleted mainly from frozen soils than in unfrozen soils because frozen conditions typically thaw of saturated and unsaturated frozen soils and implications for et al., 2000a, b). However, this comparative study identified several key cold regions’ processes with lack or poor representation in models: (1) snow redistribution and sublimation by wind including chemical transformations during transport (Pomeroy et al., 1991, 1993; Pomeroy and Jones, 1996), (2) energetics of snowmelt and areal snowcover depletion (e.g., DeBeer and Pomeroy, 2017), (3) snowpack physics (and chemistry) of flow including preferential elution (Costa et al., 2018; Costa and Pomeroy, 2019), (4) basal ice layer formation and impacts on runoff-soil interaction (Lilbæk and Pomeroy, 2008), (5) infiltration into unsaturated frozen soils including preferential infiltration of ions (Lilbæk, 2007), (6) thaw of saturated and unsaturated frozen soils and implications for erodibility and solute mixing (e.g., Edwards et al., 1995), (7) shallow subsurface flow mechanisms such as cracks and tile drainage (Zhang et al., 2016), (7) rain-on-snow and its impact on early nutrient transport, chemistry and velocity of detention flow (Costa and Pomeroy, 2019), and (8) ponding processes, depressional storage, and wetlands (Brune and Westbrook, 2012). Snow redistribution, for instance, can significantly affect the spatial patterns of snowmelt runoff generation in open windswpt agricultural regions, such as the northern US Great Plains and the Canadian Prairies in North America (e.g., Pomeroy and Li, 2000; Pomeroy and Gray, 1995). However, this process is neglected in all models, with only HYPE and SWAT accounting for snowcover heterogeneity and areal depletion via simplified parametric modelling schemes. Simultaneously, most of the models combine the temperature-index method for calculation of snowmelt with other empirical methods for computation of soil temperature, with only AnnAGNPS solving the full energy balance for both snow and soil. However, despite the simplicity of the temperature-index method being attractive for catchment simulations, research has shown that it can be problematic because of non-physical, temporally unstable and difficult to regionalize temperature index parameters (Walter et al., 2004), in addition to neglecting sublimation losses during snow ablation. This can lead to misrepresentation of snowmelt rates with consequences for the estimation of chemical transport (timing and magnitude) during this short but critical time of nutrient export (Corriveau et al., 2013). Also, the problem of infiltration into frozen soils, which may be critical for runoff, infiltration and nutrient transport forecasting in cold regions (e. g., Walter et al., 2004), can be addressed in a more physically-based manner if the dynamics of snow and soil energy balances are simulated. 4.1.4. Erosion: remains a major scientific challenge The modelling of erosion is scientifically challenging and including the impact of cold regions’ processes such as frozen soils, snowmelt intensity, and freeze-thaw cycles remains largely unrepresented in models. The empirical RUSLE method is generally regarded as the state-of-the-art with regards to the simulation of erosion at catchment scales, but it is only used in SWAT and AnnAGNPS, with the remaining models relying on other empirical formulations based on the estimation of falling raindrops and surface runoff mobilization energies. However, despite the popularity of RUSLE, this method has limitations for cold climates: (1) it is an empirical method derived from a limited set of observations, (2) it has been designed for average long-term erosion risk assessments, (3) it does not consider the effect of antecedent soil moisture conditions and soil stratification on soil cohesion, and (4) it is unsuitable for predictions of sediment transport throughout individual rainfall-runoff (or snowmelt) events because splash erosion, soil transport, and soil deposition are not treated as dynamic processes (Foster et al., 2000a, b).

In cold regions, higher erosion rates are expected over partially frozen soils than in unfrozen soils because frozen conditions typically precede unfrozen ones, and particulates are often depleted mainly from the topsoil during early snowmelt where partially frozen soils are common (e.g., Ollesch et al., 2006; Panuska and Karthikeyan, 2010). Frozen soils, which reduce soil permeability and increase runoff and often lead to significant nutrient losses during snowmelt (e.g., Ollesch et al., 2006; Panuska and Karthikeyan, 2010), are also ignored in most of the models reviewed, except for AnnAGNPS. Research shows that the relationship between the dissolved-particulate fraction and extent of soil frost is complex, varies during the melt period, and depends on soil frost dynamics. As melt starts, restricted infiltration caused by soil frost reduces soil erodibility, which results in lower particulate-P than dissolved-P losses (e.g., Tiessen et al., 2010). Freeze-thaw cycles reduce the cohesive-strength of soils increasing erodibility, which intensifies sediment transport in early snowmelt (e.g., Edwards et al., 1995). However, this central aspect of erosion in cold regions is neglected in the models.

4.1.5. Nutrient pools & biogeochemical cycles: model selection should depend on regional dominant transformation processes This section focuses on the modelling of biogeochemical processes. However, it is well known that the criteria often used for model selection is strongly based on hydrological processes. While in many cases these are needed due to lack of representation of important processes in some models (e.g., rain-on-snow is only simulated by HSPF and AnnAGNPS), the ability of models to represent biogeochemical cycling in ways that capture key regional biogeochemical processes in meaningful ways should receive careful consideration in order to maximize the outcome of the modelling effort, as well as its acceptance by different stakeholders and scientific communities (e.g., hydrologists and biogeochecists).

Most nutrient pools and biogeochemical processes are directly or indirectly simulated by all models, despite differences in the way the different mineral or organic nutrient species are grouped into shared pools (see Tables 3 and 4 for N and P, respectively). Some differences, however, may be relevant for some model applications. For example, DON and DOP are ignored in all models (except HYPE). Although this does not affect the capacity of models to capture the dominant N species exported in agricultural systems, it limits their ability to address nutrient species of growing concern, such as urea (e.g., Donald et al., 2013, 2011), which are raising growing concerns. Likewise, the simulation of DIN by some models without differentiating between NO3 and NH4 (i.e., HYPE and AnnAGNPS) hides the different controls on the fate of the different N species represented by DIN, namely nitrification, denitrification and NH4 volatilization, which are directly computed from DIN (or a hard-coded fraction of it). Finally, the processing of PON into DON and POP into DOP (all organic nutrient pools) are only computed in HYPE and HSPF, with INCA indirectly considering this process through sink terms for both NO3 and NH4. In reality, however, the equilibrium concentration of PO4 in soil solution depends on both desorption and dissolution of inorganic P, as well as mineralization of organic P (Condon et al., 2005).

Environmental factors mediate the rate of biogeochemical transformations, but there are differences in the controls characterized in the models. While the impact of temperature on N and P transformations rates is considered in all models, the effect of soil moisture is only more broadly included in the simulation of the N cycle. The impact of substrate limitation is only accounted for in HYPE for all N and P reactions via half-saturation constants. It is also considered in SWAT for denitrification and AnnAGNPS for P mineralization and desorption via fixed organic carbon concentrations or fractions. However, the temporal-spatial dynamics of such substrates may result in feedback processes via coupled cycles. For example, organic carbon may change over time in soil and riverbed sediments (e.g., seasonally) causing denitrification rates to change (e.g., Bijay-singh et al., 1988; Pfennig and McMahon, 1997), an effect that is not captured in any of the models. All models compute sorption/desorption of DOP into partP based on the sediment properties through the Freundlich isotherm or other methods. AnnAGNPS simulates transformations within mineral (partP or Mineral-P into SRP) or organic (POPres into POPact or POPstab, and POPstab
into POPact) species in greater detail than the remaining models (see Table 4), which may be more suitable for agricultural regions where the historical use of fertilizers may have lead to nutrient accumulation in soils (i.e., nutrient legacy). Here, soil temperature, soil moisture, soil pH levels, and CaCO₃ and organic carbon concentrations are selectively combined (piece-wise functions) to compute the decomposition rates. We argue that this is an essential step towards more robust conceptual models since, as opposed to the traditional basic first-order kinetic model typically used, such approaches may account for the sensitivity of soil organic matter decomposition to temperature and the quality of the substrate, which has been extensively observed (e.g., Cotrufo et al., 1994; Hartley and Ineson, 2008).

The enhancement of snowmelt nutrient release from plants subject to freeze-thaw (e.g., Liu et al., 2019a; Cober et al., 2018; Liu et al., 2013a; Macrae et al., 2010, see Section 2) is not included in any of the models, which only allow for pre-determined constant plant N and P loads (see Supplementary Material). However, recent progress in the modelling of these processes has been reported and should be considered for integration into catchment models. Costa et al. (2019a), for example, performed a series of lab experiments where alfalfa was subject to different freeze-thaw treatments, and used the results to develop the first process-based model for prediction of the temporal dynamics of N and P release from plants during snowmelt.

The coupling of biogeochemical cycles and reaction networks can lead to high complexity, a problem that is exacerbated by the mix and match of hydrological and nutrient cycles between models. While this problem does not seem to have a simple solution, we argue that the key regional processes should be identified first and used to select the simplest model possible that is able to capture the dominant processes identified.

### 4.1.6. Limitations for long-term simulations: accumulation of organic immobile pools need more attention

This review shows that while mineralization processes (i.e., from organic to mineral) are simulated in all models, synthesis/return of organic-N or organic-P (i.e., labile-PON/POP into refractory-DON/DOP, Tables 3 and 4) are only simulated in HSPF and AnnAGNPS, with HSPF being the only model to account for the sorption/immobilization of mineral-N into PON or particulate-bound-NH₄. This finding suggests that the primary emphasis of the reviewed models (as a representative sample of the current modelling capacity) is on the short-term prediction of readily available mineral-N and SRP as a response to fertilizer/manure applications. This is accomplished either through direct mineral-fertilizer inputs or through decomposition (row F, within organic N species, Table 3) followed (or directly) by mineralization of organic N and P species (row E, Table 3) originating, for instance, from manure applications. Conversely, the simulation of N and P accumulation in soils through synthesis/return of labile-PON/POP is mostly ignored, except for the most sophisticated models (AnnAGNPS and HSPF). This suggests that simpler models are particularly suited for short-term (intra-annual) simulations to capture the impact of fertilizer use and plant residue on runoff N exports, but they may be less appropriate for long-term (inter-annual) simulations where soil N and P accumulation/return and legacies may be relevant (e.g., Van Meter et al., 2016; McLauchlan, 2006). Simultaneously, the selection of the model should take into account the availability of data (e.g., type of data, data gaps and analytical issues) for model setup and verification.

### 4.2. Appropriateness of scale, scope, and complexity: further considerations for model selection (focus on biogeochemistry)

The various nutrient species are grouped into shared nutrient pools differently depending on the model, with substantial implications on the type and number of transformation pathways that can be explicitly simulated. In part, this informs the scope of the model as it dictates how complex biogeochemical processes and interactions are translated into mathematical systems. However, increasing the complexity of biogeochemical models and the number of nutrient pools has numerous challenges. More nutrient pools involve new biogeochemical pathways that need to be characterized and parameterized, a process that frequently is performed with little supporting field data. This suggests that simpler models like INCA and HYPE that require fewer parameters to calibrate may be recommended for applied research in data-poor environments. However, the nature of simpler models makes them unavoidably more deeply dependent on non-physical and unobservable parameters. This increases the reliance of these models on calibration. In turn, more detailed representation of processes in complex models like HSPF can make them arguably better suited for research studies where detailed process(es) dynamics are needed. They can also possibly be more effective as interdisciplinary collaboration tools since there is a better match between observed and modelled nutrient cycles, which may increase the perceived valued of (and trust in) the model among collaborators and help to reduce communication barriers across disciplines (e.g., modellers, hydrologists, agrologists, biogeochemists, soil scientists). However, meaningful model structures with observable and transferable parameters that match the type of primary biophysical and chemical processes occurring over cold regions catchments and scales (temporal and spatial) are needed such that catchment discretization is meaningful and used to improve model performance and integrity without excessive reliance on parameter calibration. Also, except for HSPF, the models reviewed have been primarily designed for daily time steps, which means that they cannot possibly generate hydrographs for small catchments from snowmelt events that have high temporal variability and that occur over a few days (Costa et al., 2017, 2019b). While it is true that most models can be applied at different timesteps than the daily default, and some can receive inputs at subdaily time intervals, the representation of crucial periods, where subdaily variation is important, remains inadequate. For example, the snowmelt period can have major changes in hydrology on timesteps of hours, which also affect nutrient sources and concentrations. The use of daily temperature index snowmelt models precludes any information on sub-daily melt rates. Hence, although many models use approaches such as direct linear scaling of processes and rates, key processes which change over short timescales are often not adequately represented. Examining the scalability of the simulation methods deployed in the models and shifting to approaches that minimize scale dependencies (e.g., snowpack energy balance instead of the temperature-index method for snowmelt calculations) is, therefore, a recommendation of this study.

### 5. Conclusions

The adequacy of process representations and model structural uncertainty of five popular catchment-scale hydrological-nutrient models suitable for cold regions (HYPE, INCA, SWAT, HSPF, and AnnAGNPS) has been examined to inform criteria for model selection, discuss the appropriateness of scale, scope and complexity, and provide recommendations for future research. The study involved examining the methods used for prediction of processes of general hydrology, cold regions hydrology, and N and P biogeochemical cycling. It was found that the hydrology of these models is often largely based on simplified methods and relies heavily on parameter calibration. This points at the need for meaningful model structures with observable and transferable parameters that match the type of primary biophysical and chemical processes occurring over cold regions catchments and scales (temporal and spatial). This is important to ensure that the discretization of the catchment domain is meaningful and used to improve model performance and integrity without excessive reliance on parameter calibration. The following cold regions processes were poorly represented or completely missing in most models and should be addressed in future research and model developments: (1) snow redistribution and sublimation by wind including chemical transformations during transport, (2) energetics of snowmelt and areal snowcover depletion, (3) snowpack...
Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2019.104586.