Article

Assessment of Landsat Based Deep-Learning Membership Analysis for Development of *from-to* Change Time Series in the Prairie Region of Canada from 1984 to 2018

Darren Pouliot ^{1,*}, Niloofar Alavi ¹, Scott Wilson ², Jason Duffe ¹, Jon Pasher ¹, Andrew Davidson ³ , Bahram Daneshfar ³ and Emily Lindsay ³

- ¹ Landscape Science and Technology Division, Environment and Climate Change Canada, 1125 Colonel By Drive, Ottawa, ON K1A 0H3, Canada; niloofar.alavi@Canada.ca (N.A.); jason.duffe@canada.ca (J.D.); jon.pasher@canada.ca (J.P.)
- ² Wildlife Research Division, Environment and Climate Change Canada, 1125 Colonel By Drive, Ottawa, ON K1A 0H3, Canada; scott.wilson@canada.ca
- ³ Science and Technology Branch, Agriculture and Agri-Food Canada, KW Neatby Building, 960 Carling Avenue, Ottawa, ON K1A 0C6, Canada; andrew.davidson2@canada.ca (A.D.); bahram.daneshfar@canada.ca (B.D.); emily.lindsay@canada.ca (E.L.)
- * Correspondence: darren.pouliot@canada.ca

Abstract: The prairie region of Canada is a dynamically changing landscape in relation to past and present anthropogenic activities and recent climate change. Improving our understanding of the rate, timing, and distribution of landscape change is needed to determine the impact on wildlife populations and biodiversity, ultimately leading to better-informed management regarding requirements for habitat amount and its connectedness. In this research, we assessed the viability of an approach to detect *from-to* class changes designed to be scalable to the prairie region with the capacity for local refinement. It employed a deep-learning convolutional neural network to model general land covers and examined class memberships to identify land-cover conversions. For this implementation, eight land-cover categories were derived from the Agriculture and Agri-Food Canada Annual Space-Based Crop Inventory. Change was assessed in three study areas that contained different mixes of grassland, pasture, and forest cover. Results showed that the deep-learning method produced the highest accuracy across all classes relative to an implementation of random forest that included some first-order texture measures. Overall accuracy was 4% greater with the deep-learning classifier and class accuracies were more balanced. Evaluation of change accuracy suggested good performance for many conversions such as grassland to crop, forest to crop, water to dryland covers, and most bare/developed-related changes. Changes involving pasture with grassland or cropland were more difficult to detect due to spectral confusion among classes. Similarly, conversion to forests in some cases was poorly detected due to gradual and subtle change characteristics combined with confusion between forest, shrub, and croplands. The proposed framework involved several processing steps that can be explored to enhance the thematic content and accuracy for large regional implementation. Evaluation for understanding connectivity in natural land covers and related declines in species at risk is planned for future research.

Keywords: change detection; prairie; agriculture; grassland; deep learning; membership; Landsat

Citation: Pouliot, D.; Alavi, N.; Wilson, S.; Duffe, J.; Pasher, J.; Davidson, A.; Daneshfar, B.; Lindsay, E. Assessment of Landsat Based Deep-Learning Membership Analysis for Development of *from-to* Change Time Series in the Prairie Region of Canada from 1984 to 2018. *Remote Sens.* **2021**, *13*, 634. <https://doi.org/10.3390/rs13040634>

Academic Editor: Rudiger Gens

Received: 5 January 2021

Accepted: 29 January 2021

Published: 10 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The prairie region comprising the provinces of Alberta, Saskatchewan, and Manitoba is one of the most anthropogenically modified landscapes in Canada primarily related to development and agricultural activities [1,2]. This has led to declines in natural land covers such as forest, shrub, grass, and wetland throughout or for local regions [3,4]. Climate change is also modifying the prairies landscape related to regional trends in temperature and precipitation [5]. Grasslands, in particular, are one of the least protected

and fragmented landscapes, both in Canada and globally and are an important habitat requirement for many species at risk [2,6]. In a recent study, grassland birds were found to have undergone the largest decline since 1970 relative to other breeding biomes [7] with habitat loss considered to be the primary driver of this trend [8]. Thus, an improved understanding of the rate, timing, and distribution of changes in cover types in the prairie ecosystem is needed to determine the impact on wildlife populations and biodiversity. This will lead to better-informed management regarding requirements for habitat amount and its connectedness.

In the prairie region significant advancements have been made related to crop type mapping in Canada [9,10] and the United States (US) [11]. While the Annual Space-Based Crop Inventory (ACI) produced by Agriculture and Agri-Food Canada (AAFC) provides annual crop-type mapping with high accuracy (>80%), its application to simple post-classification change detection often results in low change accuracy where the classification errors are spatially independent [12,13]. Currently it covers the period from 2009 to present. Gage et al. [14] used the AAFC and the US cropland data layers to determine plowed lands applying temporal rules to account for misclassifications. Using a similar strategy, a land-use dataset at a decadal time interval from 1990 to 2010 was developed for the prairie region by AAFC [15]. This dataset was compiled from existing land-cover data using evidence-based rules to combine and seek temporal consistency between data sources. Overall map accuracies ranged from 84% to 93%, but change accuracy between maps was not reported. Yang et al. [16] used change vector analysis with several refinement rules for a study area in Alberta, Canada to detect grassland to cultivated land conversions. Reported overall accuracy was high (89%), but producer's accuracy for the change class was lower (60%). They also found that extending the method to a larger region reduced accuracy slightly.

The development of land-cover time series specifically designed to be temporally consistent and thus suitable for map comparison can provide accurate change results capturing *from-to* class conversions. These approaches are generally based on some form of change updating, post-classification refinement, or combination of these to ensure temporal consistency [17–27]. Approaches vary in the degree of effort required to achieve good results, types of changes detected, and temporal frequency among other methodological differences. The US National Land Cover Database for example, leverages data from several national programs and fractional mapping efforts for various aspects of its development [27,28]. In Canada, land-cover time series products have mostly focused on changes associated with forests [17,20,21,26]. This is due in part to the variable nature of agriculture and grass covers making reliable detection challenging. Agriculture crops can have a diverse range of spectral responses, develop rapidly post planting in spring, senesce early in mid-to-late summer, and can be harvested multiple times a season. Grassland spectral response is often highly sensitive to moisture conditions due to shallow rooting systems [29] and is also influenced by canopy density [30]. Further, land-cover and land-use classes of interest in the prairies are often of small size relative to the spatial resolution of many daily revisit remote-sensing sensors limiting detection of urban and industrial development and mixing boundaries between crop and natural land covers. Moderate resolution sensors such as Landsat provide improved spatial resolution, but the revisit frequency and cloud cover can significantly reduce the availability of clear sky observations for analysis.

More recently, continuous change detection methods based on modeling Landsat time series have seen significant development [19,22,31,32]. These detect change occurrence within the year, can be implemented with a high degree of automation, and results can be used to provide an estimate of land cover. In these approaches, the time series is modeled and changes detected based on the model deviation to the time series across bands/features. It has been shown to be effective for more spectrally distinct and abruptly occurring changes. Subtle and gradual changes have been identified as a greater challenge, and improvements and modifications have been presented in [19,32]. For subtle changes some capacity for detection could be reduced without specific determination of

the weighting of spectral, temporal, or spatial input features. For example, if spring is the primary period to differentiate a given land cover, then including summer and fall data may decrease the relevance of the spring input reducing the ability to detect the change. However, this implies a requirement for training data to define feature weights and is a possible disadvantage regarding a more data driven automated approach. The frequency of Landsat acquisitions and associated clear-sky observations is also an important consideration [19,22,31–34].

Detection of grassland types as native and non-native is an important need in the prairie region regarding biodiversity and wildlife habitat [2]. Previous research has shown some potential to separate these grasslands and seeded forage pastures to varying degrees using seasonal data [35–37] or detecting evidence of plowing with lidar data in [38]. Results show that research is progressing, but data is currently insufficient for assessment across the entire prairie extent. Further research is also needed to evaluate the ability to detect changes between these classes over time.

Here we investigate an approach based on converting Landsat observations to class memberships using a convolutional neural network (CNN) and detecting changes in the memberships through time. Our objective was to apply and assess the viability of this approach as a means to capture *from-to* class conversions in the prairie region of Canada.

2. Materials and Methods

2.1. Overview

The main processing steps followed in this study are shown in Figure 1. This involved: (1) compositing Landsat time series for spring, summer, and fall periods from 1984 to 2018, (2) training a deep-learning classifier using the AAFC and North America Land Change Monitoring System (NALCMS) cover data, (3) applying the classifier to generate class membership time series, (4) detecting change seed points from the time series using a split window method and a set of membership tests, (5) segmenting changes using region growing from the detected seed points, and (6) removing repeat detections using a temporal cleanup procedure followed by application of a spatial-temporal mode filter. Assessment of the classification and change results was undertaken to identify the limitations and potential improvements for large extent applications.

Figure 1. Overview of processing approach.

2.2. Study Area

The study area focused on the Prairie Ecozone in the south of the Canadian provinces of Alberta, Saskatchewan, and Manitoba. Farming is the predominant human activity with much of the land utilized in some capacity for agricultural cropland, rangeland, and pasture [39]. The mean temperature varies across the region from approximately -9 to -18 °C in winter and 16 to 20 °C in summer. Annual precipitation varies from 250 mm

in the arid regions in the south-west to 700 mm in the south-east below lake Manitoba. Snowfall accounts for approximately a quarter of the precipitation [39]. Three sample regions of 90 by 150 km were selected (Figure 2). These were chosen to represent different mixes of cropland, grassland, and forest covers and associated changes.

Figure 2. Samples used for classifier training and assessment. Grey circles were used for training, yellow triangles were used for deep-learning model convergence, and red squares were used for independent accuracy assessment. Black outlines identify the areas used for change assessment. The land-cover map is the 2010 land cover of Canada [40].

2.3. Landsat Processing and Ancillary Data

Landsat-5 and -8 Collection 1 corrected to surface-reflectance level 2 data were acquired from the United States Geological Survey (USGS) and ingested into the Open Data Cube software for analysis. More than 1000 images intersected each study area. To improve separability of some classes data from spring (day of year 135 to 170), summer (180 to 235), and fall (235 to 275) were included. Only the red, near-infrared, and short-wave bands were kept to minimize residual atmosphere and haze and reduce overall processing requirements. We also included digital elevation data from the Shuttle Radar Topography Mission to improve classifier optimization across elevation strata.

Extracting consistent multi-season clear-sky Landsat observations in agriculture environments is challenging due to the diverse phenology of the cover types and irregular and sparse sampling of clear-sky Landsat observations. Thus, pooling or modeling of observations is often required. For classifier training we pooled images across three years to account for missing data and to reduce residual clouds, shadows, and haze not captured by the Landsat quality flag. Image chips of 32×32 pixels were extracted and the 40th percentile was used for compositing to provide good-quality data for training.

To apply the classifier to the time series, image composites were produced using the same method for each year, but without pooling across years. Additional outlier detection and interpolation of data for missing years was required. For outlier detection, we applied a Lowess filter and computed the z-score of the residuals. Observations with a z-score of greater than 1.35 (~80% confidence) were replaced with the smoothed estimate. Two iterations of this outlier detection procedure were applied. Missing observations in a given year were linearly interpolated as a final step.

2.4. Classifier Training and Test Samples

For classifier development, samples across the prairie region were acquired for training, verification, and final independent testing. The verification sample was used to determine training convergence of the deep-learning classifier and the test sample for independent assessment. To generate these, a systematic approach was used where samples were collected following a grid of sample blocks of 9 by 9 km with 15 km spacing between blocks. Blocks were assigned as training, validation, and test. For the test sample a column of the sample grid was randomly selected and every fourth column from that added. Only every second row for the selected columns were included in the test sample. Figure 2 shows the distribution of these sample sets. In total there were 39,520 (60%) training samples, 23,934 (30%) verification samples and 6509 (10%) test samples. Here we focused the majority of the samples for training and verification in an effort to maximize model performance for use in the change analysis. We trained the classifier over the full extent so that it could be applied across the entire region and to capture a greater range of class variability potentially improving generalization in space and time.

Training labels for these samples were derived from the AAFC ACI time series and the NALCMS circa 2010 land cover [40]. Both products were generalized to eight classes of bare, grass, pasture, crop, shrub, forest, wetland, and water following from the AAFC product specification [41]. The AAFC ACI data for 2009–2011 was used to extract the sample for Landsat 5 and 2015–2017 for Landsat 8. These samples were merged for training and testing. For each, sample pixels that were consistent across the three years were used. The NALCMS land cover was added to further check samples for natural cover types including wetland, shrub, and forest with the AAFC ACI.

After initial runs, additional training samples were acquired to improve specific classes based on interpretation of Landsat and high-resolution ESRI World Imagery. Road samples were added to better define roads, as initial results showed reduced detection where the road did not strongly contrast with the surrounding cover. Grasslands were also found to be temporally variable and additional samples were extracted from the time series. These samples were selected to be around 8 km away from a test sample to allow for capturing the class spectral and spatial variability within a region.

2.5. Classification and Assessment

In addition to the deep-learning classifier, we also tested random forest that included spatial features for comparison. Random forest is an efficient algorithm that can be easily applied in a distributed processing framework. Thus, if sufficient accuracy could be achieved then it would be the preferred classifier. For assessment, we evaluated the classification results against the test sample and report error matrices, summary accuracy measures, and class accuracies. Summary measures included the overall accuracy, kappa, and average F1-score. For class accuracies we report the F1-score, which provides class-specific accuracy, accounting for both omission and commission error.

2.5.1. Ensemble CNN

For the deep-learning method, we used a CNN incorporating multiple image extents to account for the sensitivity of CNNs to object size/scale [42–44]. It is based on the structure used in [45] and shown in Figure 3. The input features were the spring, summer, fall, and elevation data for an image chip of 32 by 32 pixels. For the CNN this image

was cropped to sizes of 1×1 , 4×4 , and 8×8 in addition to the input of 32×32 . We did not include the 16×16 size to reduce the overall size of the network. For each size, we developed a separate network path following a simple ResNet configuration, except for the 1×1 which used dense layers instead of convolution layers and no residual connections. Residual connections in CNNs were introduced by He et al. [46]. Residual connections force the next layer in the network to learn something different from the previous layers and have been shown to alleviate the problem of deep-learning models not improving performance with depth. Global average or max pooling was applied at the end of each path and dropout layers were added in an attempt to improve generalization. The outputs from each path were concatenated and fed into a set of final dense layers for class prediction. All activations were Gaussian Error Linear Units (GELU), which have shown good performance in object-recognition tasks among others [47]. Batch normalization was applied after all convolution or dense layers, except for the final output. As for ResNets, l2 weight regularization was used with a factor of 0.0001. For the final output layer after concatenation an eight-node dense layer with softmax activation was applied.

Figure 3. Schematic of the CNN used in the study. Input image size was 32 by 32 pixels shown as the large grey square. Numbers represent the number of features output from a convolution or dense layer, except for the cropping layers where the crop dimension is shown.

For training, the Adam optimizer was used with 500 epochs and categorical-cross entropy loss function. Batch size was set to 64. Data augmentation was applied to enhance sample variation and included rotation, reflectance bias of $\pm 10\%$, and random noise of $\pm 10\%$ reflectance for 15% of the input image. To determine training convergence, we monitored the validation data across training epochs and the network weights with the best performance were kept. Early stopping criteria was applied if no improvement was found in 50 epochs. For CNNs, weight initialization and random selection for batch training can lead to variability in the results. Thus, an ensemble of three models were generated and the average of the final output layer taken. Max pooling was used for two models in the ensemble and average pooling for one. In initial tests using max or average pooling showed that both provided improvement in some cases and thus could enhance the ensemble results. All CNNs were trained using the TensorFlow backend on a Tesla M100 graphic processing unit.

2.5.2. Random Forest

For random forest, input features included the mean and standard deviation for image window sizes of 5×5 , 15×15 , and 32×32 in addition to the central pixel values for each input feature. We used a random search to identify an optimal parameter set where 100 sets were sampled and compared using three-fold cross validation. Search parameter ranges were number of trees (200 to 800, by 100), minimum samples split (2–4), minimum

samples per leaf (2–4), maximum depth (full), bootstrap (true), and maximum features for split (square root of the number of features). For a description of these parameters see the Scikit-learn documentation [48]. These results were refined using a grid-search with the parameter ranges set to ± 50 for the number of estimators and ± 1 for the other parameters.

2.6. Membership Based Change Detection and Segmentation

2.6.1. Split Window Detection of Change Seed Points

To detect persistent class *from-to* changes, we generated a membership time series using the ensemble CNN applied to each year. Thus, for a given pixel we had eight class memberships for each year from 1984 to 2018. These were temporally smoothed using a filter size of three to reduce noise. To detect change for a pixel, we applied a moving split window approach and evaluated several tests based on the class memberships.

For the split window detection, a segment of the time series of size of N years was split into two $N/2$ year windows representing the *from* (year- $N/2$) and *to* (year+ $N/2$) periods for change detection. Note that the split window size ($N/2$) was truncated at the time-series edges. The following tests were applied to detect change seed points for each year in the time series for the class with the maximum median membership (maxMm) determined in the *from* or *to*-windows:

- (1) The change of the median membership in the *from*-window to the *to*-window must be greater than a change threshold (ct) for the class with the maxMm defined in the *from*-window.

$$M_{fc,y-N/2} - M_{fc,y+N/2} > ct \quad (1)$$

- (2) The change of the median membership in the *to*-window to the *from*-window must be greater than a change threshold for the class with the maxMm defined in the *to*-window. An adjustment factor for class transitions was included.

$$M_{tc,y+N/2} - M_{tc,y-N/2} > ct - adjustment \quad (2)$$

- (3) The *from-to* class memberships must occur as the maximum for a minimum occurrence threshold (ot) in each window. Here the two tests were combined.

$$\left[\frac{\sum_i^{N/2} x [m_{fc,y-i} = \max_{y-i}]}{(N/2)} \times 100 \right] + \left[\frac{\sum_i^{N/2} x [m_{tc,y+i} = \max_{y+i}]}{(N/2)} \times 100 \right] > 2ot \quad (3)$$

- (4) The median membership for a class defined in the *from* or *to*-windows must be greater than a minimum threshold (mt). Here the two tests were combined.

$$M_{fc,y-N/2} + M_{tc,y+N/2} > 2mt \quad (4)$$

- (5) Crop to bare change was detected if the crop class passed test 1, the median membership for the bare class increased by more than 15%, and the median Normalized Difference Moisture Index (NDMI) [49] in the *to*-window was less than 0.1.

$$M_{bare,y+N/2} - M_{bare,y-N/2} > 15\% \text{ and } M_{NDMI,y+N/2} < 0.1 \quad (5)$$

where $M_{fc,y-N/2}$ represents the median membership for the class with the maxMm in the *from*-window and $M_{fc,y+N/2}$ is for that class in the *to*-window. $M_{tc,y+N/2}$ represents the median membership for the class with the maxMm in the *to*-window and $M_{tc,y-N/2}$ is for that class in the *from*-window. The year is y and mfc or mtc are the membership values for the year of the class with the maxMm in the *from* or *to*-windows. The $\max$ is the maximum membership value for a given year among all classes.

After examining initial results, we added test 5 to improve detection of change from crop to bare. The crop class included cropping practices with low canopy cover and high

soil exposure. Thus, some bare areas associated with oil and gas exploration were not well-captured because the membership was confused with these low-cover crops.

For this implementation the window size was set at 24 years. The change threshold was set at 25% and no adjustment for class transition was applied. The occurrence threshold was set at 25% and the minimum membership was also set at 25%. Of these tests, the results were most dependent on 1 and 2. Figure 4 shows examples of the membership time series for several changes for different pixels with the *from*- and the *to*-windows highlighted.

Figure 4. Example membership time series for selected pixels (a–d) showing the *to* and *from* windows associated with a detected change point.

For some changes, a class such as forest may initially be bare or grass following disturbance for a short period of 1–5 years. Thus, the adjustment threshold for test 2 can be used to address this and can be set for specific change classes. However, because we used a long period of 24 years with the objective of detecting persistent class changes, we did not apply this adjustment.

2.6.2. Threshold Adjustments

To incorporate class confusion, we increased the membership change threshold for each class as a function of the commission error percent multiplied by a confusion penalty. The error percentage in its decimal form was rounded to the nearest tenths place to account for uncertainty in the estimates. The confusion penalty was set with a value of 30%. For the shrub class for example, this increased the threshold by 9% (0.3×30). For this implementation, the membership change threshold was selected to be slightly biased towards commission error. Including this adjustment increased the threshold specific to each class to reduce error.

We also adjusted the threshold at the time-series edges where the number of observations (*no*) in either the *from* or *to* windows was less than six using a natural log function ($-14.5 \times \ln(\text{no}) + 30$). However, if the number of observations was less than two the change threshold was set at 50%. This helped to reduce false detections at the edges of the time series.

2.6.3. Determining the Change Year

Once a change was detected we determined the year as the first occurrence of the *to*-class membership achieving a maximum in the *to*-window. We output the values of all tests and seed points based on test 1 to be used in spatial segmentation refinement.

2.6.4. Detection of Gradual Change Seed Points

For longer gradual changes the split window approach may miss some of these changes. To address this, we added a trend-based change detection approach applied only if change was not detected by the split window, where for each class for the full time series the Mann–Kendall trend test was applied to the membership values. If a significant trend was found and correlation using the Kendall Tau was greater 0.65 it was accepted as change. In this case the *from*-class was determined as the maximum median membership for the class from 1984 to 1989 and the *to*-class from 2013 to 2018. The change year was determined as the first occurrence of the *to*-class maximum membership in the last half of the time series.

2.6.5. Segmenting Changes

Once change seeds were detected, region growing segmentation was applied using eight-neighbor connectivity. For a given change seed, a window size of 64×64 was extracted and a mask created using a change threshold of 15% and the change type associated with the seed. This mask was segmented to delineate the change object connected with the seed point. The purpose of the segmentation was to better delineate changes.

2.6.6. Temporal Refinement

To remove repeated changes of the same type across years, a temporal cleanup procedure was applied. This occurred because the change tests can be triggered by the same change over several years. After removal of repeated occurrences, we applied a spatial–temporal mode filter of $3 \times 3 \times 3$. This helped to improve the detected change year as it was found to vary between neighboring pixels by ± 3 years.

2.7. Change Detection Assessment

To evaluate change we provided several examples of the change results relative to Landsat for various change types and carried out a more quantitative assessment to identify change classes with reduced performance. Statistically rigorous evaluation of *from*–*to* change accuracy is difficult due to the large number of change classes representing a range of abrupt, gradual, and subtle changes. For validation of a change product, good practices are recommended in [50], which recognize the importance of practical constraints such as cost and reference data availability. However, even with these considerations such accuracy assessments are a substantial undertaking and thus are not often applied in

method development and prototyping research. For this reason, we followed an approach similar to [32], involving interpretation of Landsat time series. For historical assessment such as this, Landsat interpretation can be a practical option for many change classes and can also provide an estimate of the change timing. However, for some change classes interpretation is much more ambiguous. Particularly, for small patches or for gradual subtle changes that are spectrally similar to one or more classes. To account for this in the results, we include an estimate of interpreter confidence.

For this assessment, we sampled five changes for each change class where possible and ten samples for each no-change class for each of the study sites. For the no-change samples, the 2018 AAFC ACI product was used for sample stratification. For each sample we extracted the Landsat scenes for the point for 60×60 , 20×20 , and 3×3 sized windows and examined this multi-extent image time series to determine if there was evidence to support the detected *from-to* change occurrence. Where suitable, high-resolution ESRI World Imagery was consulted. The year of the change was also estimated. For the no-change samples, we examined the time series to determine if a change had been missed.

To improve understanding of changes associated with the pasture class we split it into high and low biomass pasture. The high-biomass pasture class represented more intensively managed or seeded-forage pastures. In some cases, it also included mixes of pasture and shrub or naturally occurring high-biomass grasslands typically associated with wetland areas. The low-biomass pasture class represented a grassland- or low-vegetation-like cover condition. To separate these, we used the class memberships and NDMI, where if the membership for crop was greater than the membership for grassland by 5% or if the NDMI was greater than 0.18 it was identified as high-biomass pasture.

Including no-change there were a total of 81 classes, which makes standard accuracy assessment impractical to implement and present. Furthermore, not all change classes will be of interest nor can all be assessed with the same confidence. For these reasons we compute the estimated percent correct of each change class and provide it as a 9 by 9 matrix for convenience. The diagonal represents the percent correct for no-change classes, and the off-diagonal the change classes.

3. Results and Discussion

3.1. Classification

The deep-learning ensemble CNN achieved a higher accuracy than random forest by 6% for the average summary F1-score. More importantly, class accuracies were more balanced (Table 1). Classes where spatial properties were evident increased the most. For example, the CNN was better able to separate croplands with high soil exposure from other bare land covers. All classes except for wetland and shrub had an F1-score greater than 80%. Shrub had the lowest class accuracy and was confused with several classes, but most predominantly with wetland and forest (Table 2). Wetland was mostly confused with forest, followed by crops. These results are generally consistent with other large-extent land-cover mapping applications using Landsat in Canada [15,26,40], the US [51], and with the reported accuracies in the AAFC ACI metadata [41].

Table 1. Comparison of the ensemble CNN and random forest classifiers.

Classifier	Overall			F1-Score							
	Accuracy	Kappa	F1_Ave	Bare	Grass	Pasture	Water	Crop	Forest	Wetland	Shrub
Ensemble CNN	0.884	0.861	0.840	0.826	0.867	0.835	0.978	0.929	0.920	0.763	0.600
Random Forest	0.845	0.816	0.776	0.650	0.842	0.771	0.981	0.901	0.890	0.722	0.450

Table 2. Error matrix for the ensemble CNN.

		Reference								R_Sum	UA
		Bare	Grass	Pasture	Water	Crop	Forest	Wetland	Shrub		
Predicted	Bare	145	4	8	0	15	3	4	1	180	0.81
	Grass	5	828	92	0	16	5	21	10	977	0.85
	Pasture	8	92	1014	2	56	17	19	22	1230	0.82
	Water	0	1	3	887	1	1	18	1	912	0.97
	Crop	11	3	49	1	1343	0	27	1	1435	0.94
	Forest	1	2	18	2	7	1045	64	32	1171	0.89
	Wetland	0	0	6	9	14	20	391	30	470	0.83
	Shrub	1	3	8	0	3	9	11	99	134	0.74
	C_Sum	171	933	1198	901	1455	1100	555	196		
PA		0.85	0.89	0.85	0.98	0.92	0.95	0.70	0.51		

Note: C_Sum—column sum, R_Sum—row sum, PA—producer's accuracy, UA—user's accuracy.

3.2. Change Detection

Assessment of the change results are shown in Figure 5 as matrices of *from-to* change classes providing the estimated agreement with reference samples, the number of samples used to evaluate each change class, and the error in the change year. These are combined across all sample regions and noteworthy differences between regions are discussed below. Examples for visual evaluation are provided in Figure 6. These reveal the delineation accuracy of the change objects and help provide confidence with the derived results.

Figure 5. Change assessment by change class (a) percent agreement with reference samples, (b) number of reference samples, (c) mean absolute difference in the change year.

Figure 6. Example change detection results; (a) grass to crop and road, (b) crop to urban, (c) water-body shrinkage, (d) forest to crop, (e) crop to pasture, and (f) pasture to forest. The order from left to right is the result for the change class and year followed by example Landsat images displayed as the near-infrared, short-wave infrared, and red band combination.

Reasonable detection performance was found for many of the change classes. Classes with distinct spectral properties and persistent cover conversions were better detected, as expected. Detection errors were associated with class spectral confusion, small change objects, infrequent changes, changes close to the start or end of the time series, and temporal variability of the land-cover classes. These factors are generally known to affect classification and change-detection performance, but the specific manner in which they influenced the results of this approach are worth highlighting. Spectral confusion in this case caused an overall reduction in the derived memberships and increased the membership variability between years, reducing detection of some changes. It also caused some commission error, where for a given area the occurrence of spectrally similar classes increased over time, leading to false detection. This was most problematic for samples located near object boundaries. For example, shrub and crops can have similar spectral properties. Thus, a small patch of shrub surrounded initially by grass that is converted to a crop can cause this type of error. This is in part why smaller, less spectrally distinct changes were more difficult to detect. Furthermore, smaller objects do not benefit from the ability of the CNN to represent spatial features. Detecting changes at the start or end of the time series can be more challenging because of the reduced number observations available resulting in greater sensitivity to classification error. For some classes, temporal variability on an annual time scale presented a challenge. For example, depending on climate in a given year, wetland areas can appear as water or dryland. In this case, determining if change occurred has a higher degree of uncertainty and is more dependent on the specific definition of what constitutes a change in terms of temporal persistence.

Computing a summary accuracy measure from these results could be potentially biased due to ambiguity in the interpretation of some change classes related to spectral confusion and inadequate spatial or temporal differences. Furthermore, effectively representing omission error can be problematic because its occurrence is rare relative to the sample space (area of no-change) reducing the likelihood of sampling a missed change. However, recognizing these limitations, these results can be collapsed to change and no-change classes giving an estimated overall accuracy of 85% and average F1-score of 83%. Comparison to other studies is problematic because of substantial differences in study areas, change classes, and assessment methods.

The timing of the detected changes was good on average ± 2 years. However, large errors did occur related to gradual changes, *from-to* classes with a high degree of temporal variability, or where the conversion included a period of a change transition class. Errors of more than 8 years occurred for less than 1% of the sample and were never larger than 10 years. Part of the error in this case was related to the difficulty of precisely interpreting change timing for gradual changes.

The percent change relative to the total area of the study sites is shown in Figure 7. Values were multiplied by 1000 to better represent small changes, where 10,000 equals 10%. Two graphs with different y-axis ranges are provided for visualization of less frequent changes. Considering the challenge of describing the results for all change classes, we elected to summarize these for natural covers of loss and gain for grassland, forest, water/wetland combined, and also highlight some of the other significant notable changes observed.

Figure 7. Percent change detected by change class across study sites for the full y-axis range in (a) and a smaller range in (b). The wetland class is denoted as *WetL* and the high-biomass and low-biomass pasture classes as *PastH* and *PastL* respectively.

3.2.1. Grassland Loss and Gain

The primary cause of grassland loss was due to conversion to some form of agriculture (pasture and crop in Figure 7). It occurred mainly in the south-west study region. Conversion to cropland and high-biomass pasture showed good agreement with reference samples and is supported by qualitative assessment of the results. Figure 6a shows an example of grassland converted to pivot-irrigated cropland. Results for the larger region are shown in Figure 8. Conversions between grassland and low-biomass pasture were less accurate due to spectral and spatial similarity. Interpretation confidence was low and also a factor in this result. Conversion to bare (access road) is also evident in Figure 6a and was another source of grassland loss, albeit small in terms of total area. Some grassland was lost due to increases in water and wetlands discussed further in Section 3.2.3.

Gains in grasslands were predominately from agriculture. These do not represent a gain in natural grassland. In this case the conversion represents an area where agriculture has not been practiced on the land for several years and may contain a mix of native and non-native species. Similar to the strategy applied by [38], these conversions can be used to distinguish native and non-native grasslands based on evidence of previous agricultural activity. However, specific to this approach, areas of non-native grassland that had been cultivated prior to 1984 would be missed. Seasonal differences in Landsat imagery have been used for native and non-native grassland detection in [35,36] and could be explored using this framework.

Figure 8. Grassland loss and gain in the south-west study region from 1984 to 2018.

3.2.2. Forest Loss and Gain

Forest loss was caused by conversion primarily to agriculture followed by water/wetland and then developed land. Figure 6d shows an example of forest conversion to cropland. Forest conversion to water/wetland occurred due to the development of artificial water bodies or natural accumulation of water in low-lying areas. In some cases, it was related to mixed pixel effects, where an increase in water caused it to become the dominant cover in the pixel, where previously it was present but not dominant. Forest to shrub occurred in a few small areas, but due to forest–shrub confusion it had a higher degree of commission

error. In most cases, it captured the impact of forest canopy damage in a given year, but was not considered to be shrub cover.

Regarding forest gain, the primary conversion was from agriculture, with high-biomass pasture areas being the primary contributor. Figure 6f shows an example. Most of the forest gain occurred in the north-west study region. Conversion from crop to forest was incorrectly detected in some cases because of small forest object size, spectral confusion, gradual change rate, and the use of the CNN which can be influenced by the surrounding land cover [45]. Pasture to forest conversion in the south-east study region was incorrectly detected. It occurred mostly between 1987 and 1990. Evidence of drought and insect defoliation occurred prior to this period and the detected change was likely related to forest recovery [52,53]. Applying a higher threshold to the change membership can be used to reduce or remove this error in this region. In the north-west region pasture to forest conversion was more accurately detected. There was little forest in the south-west region and thus little to no forest gain or loss was detected.

3.2.3. Water/Wetland Loss and Gain

Trends in landscape moisture have increased for parts of the southern prairie region [5,54] resulting in more water and wetlands. In these cases, wetlands have transitioned to a more persistent water state and similarly cropland/pastures have transitioned to a wetland state. In some areas classes may fluctuate between water, wetland, and dryland classes interannually. Temporal persistence, in effect, ignores these short-term fluctuations so that the dominant class state can be represented in the change time series. Other approaches have been developed for annual water mapping with Landsat [55].

For the southern study regions, water loss occurred mostly due to conversion to wetlands associated with localized changes in wetland hydrology for a period of several years. Gains in water were due to conversions mostly from wetlands, agriculture classes, and some bare areas. Gains in wetland occurred as conversions from agriculture classes with some from bare areas and grasslands. In the north-western study region, loss and gains were more locally variable compared to the southern areas. For water, there was an overall reduction in the north-west. This was caused by water bodies drying out to become wetlands and bare areas (Figure 6c). Wetlands increased in the area due to the conversion from water, but again it was locally variable with some wetland loss caused by conversion to water. In these cases, mixed pixels were also a factor.

3.2.4. Change in Cropping Practices

The most frequent changes detected in all study regions were related to cropping practices, mostly conversions with pasture and cropland. The pasture class consisted of areas cleared for grazing, previously used for some form of agriculture production or currently cultivated non-native grasses used for feed or seed production. An important distinction between cropland and cultivated pasture was based on previous findings that these pastures were distinguishable from grasslands and croplands related to earlier spring greenup [37]. However, this has not been widely examined across the prairie region.

In the south-west study region, another conversion of interest was a change from low-canopy-cover to high-canopy-cover cropping practices. This was detected as a conversion from bare to cropland, but due to confusion between bare areas and low-canopy-cover croplands, omission error is expected to be high. The opposite conversion of cropland to bare was also caused by changes in cropping practices, but in some cases it was due to a change to pivot irrigation leaving field edges and corners in either a plowed or low vegetation cover state.

3.2.5. Roads

At 30 m spatial resolution the detection of roads is challenging due to the width of the road and spectral contrast with the surrounding cover. For change detection these factors add additional complexity as they can change over time. Examining the results,

new roads in forest and shrub tended to be captured because of sufficient contrast between the forest and road. In grasslands many roads were detected, but small roads and trails were missed and in general road detection in grasslands was more difficult due to reduced contrast relative to forests. Roads in cropland areas were the most problematic and were a source of commission error. For example, consider a time series where cropping practices change from low-canopy-cover to high-canopy-cover crops. In this case, the road is difficult to detect in the low-canopy condition because the road and adjacent low-canopy-cover cropland have similar spectral properties. In the high-canopy-cover cropland, the road is more easily detected, causing commission error in some cases. Improving training of roads and road edges would help to correct this, but also examination of the temporal trend in the surrounding land cover could be used to identify cases of high road-commission error associated with croplands.

3.3. Considerations for Large-Extent Application

Accurately capturing *from-to* class change is challenging depending on the class set, where more classes tend to reduce separability of spectral, spatial, and temporal properties, ultimately leading to reduced change-detection accuracy. Furthermore, changes can occur over a range of temporal scales and be highly temporally dynamic, taking several years to reach a constant state. Results of the approach investigated revealed good detection performance for many change classes and several improvements were identified to be considered for large-extent application and additional research. In the following, we highlight some of these more important considerations involving Landsat processing, class memberships, and change detection.

3.3.1. Landsat Processing

In this research, we used Landsat Collection 1 data; however, significant improvements have been made for Collection 2, released in late 2020, regarding geolocation among other factors [56]. Geolocation is a critical aspect for change detection and any improvements should translate to better change detection, particularly for small-change objects [57,58]. Improved radiometric correction is also a benefit, but is less of a concern because the CNN was trained using a data augmentation strategy that varied the reflectance within $\pm 10\%$. Furthermore, spatial and temporal features are expected to help reduce sensitivity to these errors.

For the seasonal composite generation there are several approaches that could be used. Time-series modeling methods suitable for Landsat [22,31,32,34,59] could provide an improvement by better representing seasonal properties, particularly in spring during the short period of vegetation canopy development. However, due to sparse sampling of good-quality observations caused by Landsat revisit frequency and cloud and cloud shadow, it was unclear if a significant improvement would be achieved. For this reason, we elected to devise a more computationally efficient strategy and postulated that the spectral-spatial features captured by the CNN and the evidence-based change detection approach could compensate for some data limitations. Investigating the integration of time-series modeling approaches is planned for future development.

3.3.2. Class Memberships

Classes selected and how they are defined is important to consider. The greater the spectral confusion, the more error is likely to occur in the change analysis. Spatial properties are useful for larger landscape objects. However, for small objects, spectral/temporal properties are the dominant discriminators. At the Landsat scale and for the prairie region in Canada, which is considered a fragmented landscape [60], small objects were more common for classes such as forests, shrubs, wetlands, and developed lands. Improving small-object detection needs to be further examined and may require the inclusion of additional input features or higher weighting of the single-pixel network in the CNN

ensemble. For many land-cover products a minimum mapping unit of several pixels is considered in production or validation recognizing this limitation [23,61,62].

The sensitivity of the change accuracy attributable to classification error was also hypothesized to be reduced by the examination of persistent change evidence in the time-series memberships. This, however, turned out to be only partly true as class confusion for a given pixel was sometimes biased as opposed to random over the time series. Sensitivity to random class confusion was less problematic because temporal persistence criteria can reduce its effect. There are numerous potential enhancements regarding classification that could lead to better-defined class memberships used in the change detection. As with all classification methods improvements in training data is expected to enhance the results. In this analysis, a large spatially extensive sample was generated, but its quality in space and time were variable. This led to membership confusion in some regions and time periods, resulting in change error. For this analysis, the assumption that a deep-learning global model could perform sufficiently within this change framework proved to be a limitation on the change quality. However, this provided a large training dataset to improve model generalization over space and time. Locally retraining this model with a small high-quality sample can be used to improve the results for specific strata following a transfer learning approach. This strategy, to learn from the existing AAFC data, was used to leverage investment and knowledge in previous efforts and reduce sampling requirements for local refinement.

There are numerous aspects to consider for enhancing the deep-learning classifier, specifically regarding the input image size; network configuration, depth, and width; and training strategy. In initial experimentation, several standard forms of deep-learning CNN architectures were tested based largely on ResNet [46] and Inception [63], but these did not perform well due to the range of image objects size/scales encountered. To account for objects represented as a single pixel up to larger assemblages extending beyond the 32 by 32 pixel input image, the ensemble with varying input image sizes was selected. Segmentation-based approaches were also considered as an alternative, but sufficient quality training data for segmentation was a limitation.

3.3.3. Change Detection

For detecting changes in the membership time series, we decided to use a set of tests that were considered intuitive to understand and thus define parameters for. However, comparison of various time-series change-point methods could be evaluated for this purpose. A detailed review of methods is given in [64]. We did explore the use of a *t*-test for detecting change points. However, using a statistical criteria such as 95% led to a large number of commission errors and thus selecting a suitable *t*-value compared to the membership change threshold we felt was less clear.

Class specific tests or thresholds can be defined to improve the results regarding the membership change and temporal class persistence. The size of the split window can also be varied to capture changes occurring over different temporal scales. For example, conversions between water and other cover types can occur annually and can be detected using smaller temporal windows, whereas conversion to forest needs to be considered over a longer period. This temporal membership analyses approach provides flexibility in defining criteria to capture change which is an advantage, but also adds complexity.

For some changes a short transitional class occurred such as the case of lake expansion where an intermediate wetland condition may arise before becoming water. The current change implementation does not detect these transitions as they tend to be short occurrences of 2–4 years, but will detect the start and end class states. Refined tests can be developed to detect these transitions if desired. For this implementation, the 24-year split window detection approach captured most of the change, except for some changes related to water and pasture or croplands that occurred on shorter time scales. The long-term change component did not detect a lot of seed pixels and likely could be removed for this implementation.

4. Conclusions

Here we applied a deep-learning temporal membership approach to detect land-cover changes in the prairie region of Canada. Examination of class memberships provides a framework that can be used to examine evidence of change and potentially allows for more optimized weighting of spectral, temporal, or spatial features to improve subtle change detection. The approach is flexible and can be adapted to detect a large range of *from-to* class changes over varying temporal scales. In the prairie region, change detection is challenging because landscape objects are often small represented by a few pixels in moderate spatial resolution imagery. Furthermore, for many classes there is strong spectral variability between years, related to climate conditions and land-use activities such as crop selection, cultivation practices, harvesting, and grazing. The results of this approach show that many *from-to* class changes were well-captured. Following from this assessment several improvements are suggested for large-extent application. Expanding on this across the prairies and examining impacts on wildlife and connectivity is planned for future research.

Author Contributions: Conceptualization, D.P., S.W., J.D., J.P., A.D., B.D. and E.L.; data curation, D.P., N.A. and E.L.; investigation, D.P., N.A. and S.W.; methodology, D.P., N.A., S.W., A.D., B.D. and E.L.; project administration, J.D. and J.P.; resources, J.D. and J.P.; software, D.P.; validation, N.A.; writing—original draft, D.P. and N.A.; writing—review & editing, D.P., N.A., S.W., J.D., J.P., A.D., B.D. and E.L. All authors have read and agreed to the published version of the manuscript.

Funding: Research was supported through a Canadian Space Agency grant for the project “Integrated Earth Observation Monitoring for Essential Ecosystem Information: Resilience to Ecosystem Stress and Climate Change”.

Acknowledgments: We would like to thank Charles Francis for his input and support of this research and Daniel Bazor for his contributions to the development of Landsat processing. We would also like to recognize Rasim Latifovic as conceptualization of the approach was influenced by discussions on the topic.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Riley, J.L.; Green, S.E.; Brodribb, K.E. *A Conservation Blueprint for Canada's Prairies and Parklands*; Nature Conservancy Canada: Toronto, ON, Canada, 2007.
2. Federal, Provincial, and Territorial Governments of Canada. *Canadian Biodiversity: Ecosystem Status and Trends 2010 Report*; Canadian Councils of Resource Ministers: Ottawa, ON, Canada; Federal, Provincial, and Territorial Governments of Canada: Ottawa, ON, Canada, 2010.
3. Canadian Wildlife Service. *The Federal Policy on Wetland Conservation*; Government of Canada: Ottawa, ON, Canada, 1991.
4. Watmough, M.D.; Schmoll, M.J. *Environment Canada's Prairie & Northern Region Habitat Monitoring Program Phase II: Recent Habitat Trends in the Prairie Habitat Joint Venture*; Environment Canada; Canadian Wildlife Service: Edmonton, AB, Canada, 2007.
5. Zhang, X.; Flato, G.; Kirchmeier-Young, M.; Vincent, L.; Wan, H.; Wang, X.; Rong, R.; Fyfe, J.; Li, G.; Kharin, V.V. *Canada's Changing Climate Report: Temperature and Precipitation Across Canada*; Bush, E., Lemmen, D., Eds.; Government of Canada: Ottawa, ON, Canada, 2019; pp. 112–193.
6. Hoekstra, J.M.; Boucher, T.M.; Ricketts, T.H.; Roberts, C. Confronting a biome crisis: Global disparities of habitat loss and protection. *Ecol. Lett.* **2005**, *8*, 23–29. [[CrossRef](#)]
7. Rosenberg, K.V.; Dokter, A.M.; Blancher, P.J.; Sauer, J.R.; Smith, A.C.; Smith, P.A.; Stanton, J.C.; Panjabi, A.; Helft, L.; Parr, M.; et al. Decline of the North American avifauna. *Science* **2019**, *366*, 120–124. [[CrossRef](#)] [[PubMed](#)]
8. North American Bird Conservation Initiative. *The State of North America's Birds 2016*; Environment and Climate Change Canada: Ottawa, ON, Canada, 2016; p. 8.
9. Fiset, T.; Davidson, A.; Daneshfar, B.; Rollin, P.; Aly, Z.; Campbell, L. Annual space-based crop inventory for Canada: 2009–2014. *Int. Geosci. Remote Sens. Symp.* **2014**, 5095–5098. [[CrossRef](#)]
10. Davidson, A.M.; Fiset, T.; McNairn, H.; Daneshfar, B. Detailed crop mapping using remote sensing data (crop data layers). In *Handbook on Remote Sensing for Agricultural Statistics (Chapter 4)*; Delincé, J., Ed.; GSARS: Rome, Italy, 2017; pp. 91–117.
11. Boryan, C.; Yang, Z.; Mueller, R.; Craig, M. Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto Int.* **2011**, *26*, 341–358. [[CrossRef](#)]
12. Stow, D.A.; Tinney, L.R.; Estes, J.E. Deriving land use/land cover change statistics from LANDSAT: A study of prime agricultural land. *Proc. Int. Symp. Remote Sens. Environ.* **1980**, *2*, 1227–1237.
13. Van Oort, P.A.J. Interpreting the change detection error matrix. *Remote Sens. Environ.* **2007**, *108*, 1–8. [[CrossRef](#)]

14. Gage, A.M.; Olimb, S.K.; Nelson, J. Plowprint: Tracking cumulative cropland expansion to target grassland conservation. *Gt. Plains Res.* **2016**, *26*, 107–116. [\[CrossRef\]](#)
15. Agriculture and Agri-food Canada Land Use 1999, 2000 & 2010. Available online: <https://open.canada.ca/data/en/dataset/18e3ef1a-497c-40c6-8326-aac1a34a0dec> (accessed on 15 July 2019).
16. Yang, X.; Smith, A.M.; Hill, M.J. Updating the Grassland Vegetation Inventory using change vector analysis and functionally-based vegetation indices. *Can. J. Remote Sens.* **2017**, *43*, 62–78. [\[CrossRef\]](#)
17. Latifovic, R.; Pouliot, D. Multitemporal land cover mapping for Canada: Methodology and products. *Can. J. Remote Sens.* **2005**, *31*, 347–363. [\[CrossRef\]](#)
18. Xian, G.; Homer, C.; Fry, J. Updating the 2001 National Land Cover Database land cover classification to 2006 by using Landsat imagery change detection methods. *Remote Sens. Environ.* **2009**, *113*, 1133–1147. [\[CrossRef\]](#)
19. Brown, J.F.; Tollerud, H.J.; Barber, C.P.; Zhou, Q.; Dwyer, J.L.; Vogelmann, J.E.; Loveland, T.R.; Woodcock, C.E.; Stehman, S.V.; Zhu, Z.; et al. Lessons learned implementing an operational continuous United States national land change monitoring capability: The Land Change Monitoring, Assessment, and Projection (LCMAP) approach. *Remote Sens. Environ.* **2020**, *238*, 111356. [\[CrossRef\]](#)
20. Pouliot, D.; Latifovic, R.; Zabcic, N.; Guindon, L.; Olthof, I. Development and assessment of a 250 m spatial resolution MODIS annual land cover time series (2000–2011) for the forest region of Canada derived from change-based updating. *Remote Sens. Environ.* **2014**, *140*, 731–743. [\[CrossRef\]](#)
21. Colditz, R.R.; Pouliot, D.; Llamas, R.M.; Homer, C.; Latifovic, R.; Ressler, R.A.; Tovar, C.M.; Hernández, A.V.; Richardson, K. Detection of North American land cover change between 2005 and 2010 with 250 m MODIS Data. *Photogramm. Eng. Remote Sensing* **2014**, *80*, 918–924.
22. Zhu, Z.; Woodcock, C.E. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sens. Environ.* **2014**, *144*, 152–171. [\[CrossRef\]](#)
23. Homer, C.C.G.; Dewitz, J.J.A.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.N.D.; Wickham, J.D.J.; Megown, K. Completion of the 2011 National Land Cover Database for the conterminous United States representing a decade of land cover change information. *Photogramm. Eng. Remote Sens.* **2015**, *81*, 345–354. [\[CrossRef\]](#)
24. Pouliot, D.; Latifovic, R. Land change attribution based on Landsat time series and integration of ancillary disturbance data in the Athabasca Oil Sands region of Canada. *GIScience Remote Sens.* **2016**, *53*, 382–401. [\[CrossRef\]](#)
25. ESA Land Cover CCI Product User Guide Version 2. Tech. Rep. Available online: http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf (accessed on 5 January 2020).
26. Hermosilla, T.; Wulder, M.A.; White, J.C.; Coops, N.C.; Hobart, G.W. Disturbance-informed annual land cover classification maps of Canada's forested ecosystems for a 29-year Landsat time series. *Can. J. Remote Sens.* **2018**, *44*, 67–87. [\[CrossRef\]](#)
27. Jin, S.; Homer, C.; Yang, L.; Danielson, P.; Dewitz, J.; Li, C.; Zhu, Z.; Xian, G.; Howard, D. Overall methodology design for the United States national land cover database 2016 products. *Remote Sens.* **2019**, *11*, 2971. [\[CrossRef\]](#)
28. Homer, C.; Dewitz, J.; Jin, S.; Xian, G.; Costello, C.; Danielson, P.; Gass, L.; Funk, M.; Wickham, J.; Stehman, S.; et al. Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land Cover Database. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 184–199. [\[CrossRef\]](#)
29. Weaver, J.E. Classification of root systems of forbs of grassland and a consideration of their significance. *Ecology* **1958**, *39*, 393–401. [\[CrossRef\]](#)
30. Guo, X.; Zhang, C.; Wilmshurst, J.; Sissons, R. Monitoring grassland health with remote sensing approaches. In *Prairie Perspectives, Geographical Essays*; Noble, B.F., Martz, D.J.F., Aitken, A.E., Eds.; University of Saskatchewan: Saskatoon, SK, Canada, 2005.
31. Zhu, Z.; Woodcock, C.E.; Olofsson, P. Continuous monitoring of forest disturbance using all available Landsat imagery. *Remote Sens. Environ.* **2012**, *122*, 75–91. [\[CrossRef\]](#)
32. Zhu, Z.; Zhang, J.; Yang, Z.; Aljaddani, A.H.; Cohen, W.B.; Qiu, S.; Zhou, C. Continuous monitoring of land disturbance based on Landsat time series. *Remote Sens. Environ.* **2020**, *238*, 111116. [\[CrossRef\]](#)
33. Kovalsky, V.; Roy, D.P. The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30m Landsat data product generation. *Remote Sens. Environ.* **2013**, *130*, 280–293. [\[CrossRef\]](#)
34. Pouliot, D.; Latifovic, R. Reconstruction of Landsat time series in the presence of irregular and sparse observations: Development and assessment in north-eastern Alberta, Canada. *Remote Sens. Environ.* **2018**, *204*, 979–996. [\[CrossRef\]](#)
35. McInnes, W.S.; Smith, B.; McDermid, G.J. Discriminating native and nonnative grasses in the dry mixedgrass prairie with MODIS NDVI time series. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 1395–1403. [\[CrossRef\]](#)
36. Olimb, S.K.; Dixon, A.P.; Dolfi, E.; Engstrom, R.; Anderson, K. Prairie or planted? Using time-series NDVI to determine grassland characteristics in Montana. *Geojournal* **2018**. [\[CrossRef\]](#)
37. Lindsay, E.J.; King, D.J.; Davidson, A.M.; Daneshfar, B. Canadian prairie rangeland and seeded forage classification using multiseason Landsat 8 and summer RADARSAT-2. *Rangel. Ecol. Manag.* **2018**. [\[CrossRef\]](#)
38. Fisher, R.J.; Sawa, B.; Prieto, B. A novel technique using LiDAR to identify native-dominated and tame-dominated grasslands in Canada. *Remote Sens. Environ.* **2018**. [\[CrossRef\]](#)
39. Ecological Framework of Canada. Available online: <http://ecozones.ca> (accessed on 15 May 2019).
40. Latifovic, R.; Pouliot, D.; Olthof, I. Circa 2010 Land Cover of Canada: Local optimization methodology and product development. *Remote Sens.* **2017**, *9*, 1098. [\[CrossRef\]](#)

41. Agriculture and Agri-food Canada ISO 19131 Annual Crop Inventory—Data Product Specifications. Available online: <http://130.179.67.140/dataset/f2f37ad5-0afa-4fc4-885d-6a6e9f8b6dc8/resource/04a7532d-a7b0-416d-939a-fb950ca8012e/download/iso19131aafcannualcropinventorydataproducspecifications.pdf> (accessed on 10 April 2019).
42. Sermanet, P.; Eigen, D.; Zhang, X.; Mathieu, M.; Fergus, R.; LeCun, Y. OverFeat: Integrated recognition, localization and detection using convolutional networks. *arXiv* **2013**, arXiv:1312.6229.
43. Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 11–18 December 2015; pp. 1440–1448.
44. Hu, X.; Xu, X.; Xiao, Y.; Chen, H.; He, S.; Qin, J.; Heng, P.-A. SINet: A scale-insensitive convolutional neural network for fast vehicle detection. *IEEE Trans. Intell. Transp. Syst.* **2018**, 1–10. [\[CrossRef\]](#)
45. Pouliot, D.; Latifovic, R.; Pasher, J.; Duffe, J. Assessment of convolution neural networks for wetland mapping with Landsat in the central Canadian boreal forest region. *Remote Sens.* **2019**, *11*, 772. [\[CrossRef\]](#)
46. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. *arXiv* **2015**, arXiv:1512.03385.
47. Hendrycks, D.; Gimpel, K. Gaussian Error Linear Units (GELUs). *arXiv* **2016**, arXiv:1606.08415.
48. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
49. Gao, B. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [\[CrossRef\]](#)
50. Olofsson, P.; Foody, G.; Martin, H.; Stephen, V.S.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2013**, *148*, 76. [\[CrossRef\]](#)
51. Wickham, J.; Stehman, S.V.; Gass, L.; Dewitz, J.A.; Sorenson, D.G.; Granneman, B.J.; Poss, R.V.; Baer, L.A. Thematic accuracy assessment of the 2011 National Land Cover Database (NLCD). *Remote Sens. Environ.* **2017**, *191*, 328–341. [\[CrossRef\]](#)
52. Sutton, A.; Tardif, J.C. Dendrochronological reconstruction of forest tent caterpillar outbreaks in time and space, western Manitoba, Canada. *Can. J. For. Res.* **2007**, *37*, 1643–1657. [\[CrossRef\]](#)
53. Bonsal, B.R.; Wheaton, E.E.; Chipanshi, A.C.; Lin, C.; Sauchyn, D.J.; Wen, L. Drought research in Canada: A review. *Atmos. Ocean* **2011**, *49*, 303–319. [\[CrossRef\]](#)
54. Luthcke, S.B.; Sabaka, T.J.; Loomis, B.D.; Arendt, A.A.; McCarthy, J.J.; Camp, J. Antarctica, Greenland and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution. *J. Glaciol.* **2013**, *59*, 613–631. [\[CrossRef\]](#)
55. Pekel, J.F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [\[CrossRef\]](#) [\[PubMed\]](#)
56. United States Geological Survey. Landsat Collection 1 vs. Collection 2 Summary. Available online: <https://www.usgs.gov/media/files/landsat-collection-1-vs-collection-2-summary> (accessed on 5 December 2020).
57. Dai, X.; Khorram, S. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1566–1577. [\[CrossRef\]](#)
58. Roy, D.P. The impact of misregistration upon composited wide field of view satellite data and implications for change detection. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 2017–2032. [\[CrossRef\]](#)
59. Zhu, Z.; Woodcock, C.E.; Holden, C.; Yang, Z. Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. *Remote Sens. Environ.* **2015**, *162*, 67–83. [\[CrossRef\]](#)
60. Roch, L.; Jaeger, J.A.G. Monitoring an ecosystem at risk: What is the degree of grassland fragmentation in the Canadian Prairies? *Environ. Monit. Assess.* **2014**, *186*, 2505–2534. [\[CrossRef\]](#)
61. Latifovic, R.; Homer, C.; Ressler, R.; Pouliot, D.; Hossain, S.N.; Colditz, R.R.; Olthof, I.; Giri, C.P.; Victoria, A. North American Land-Change Monitoring System. In *Remote Sensing of Land Use and Land Cover: Principles and Applications*; Giri, C.P., Ed.; CRC Press: Boca Raton, FL, USA; Taylor & Francis: Boca Raton, FL, USA, 2016; pp. 303–323. ISBN 9781420070750.
62. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2014**, *103*, 7–27. [\[CrossRef\]](#)
63. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **2015**, 1–9. [\[CrossRef\]](#)
64. Truong, C.; Oudre, L.; Vayatis, N. Selective review of offline change point detection methods. *Signal Process.* **2020**, *167*, 107299. [\[CrossRef\]](#)