3A.3 USING CANADIAN GEM OUTPUT FOR FORECASTS OF THUNDERSTORM INITIATION ON THE CANADIAN PRAIRIES

Neil M. Taylor[1*] and William R. Burrows[2]

[1]Hydrometeorology and Arctic Lab, Environment Canada, Edmonton, AB
[2]Cloud Physics and Severe Weather Research Section / Hydrometeorology and Arctic Lab, Environment Canada, Edmonton, AB

1. INTRODUCTION

Predicting the location and timing of thunderstorm initiation[**] (TI) is critical for operational forecasters to issue timely and accurate severe weather watches and warnings. The basic requirements for TI are sufficient moisture, instability, and lift for air parcels to reach their level of free convection (LFC) and maintain positive buoyancy through the troposphere. This, or course, is a very simplified conceptual model. The potential for TI is sensitive to complicated processes both through the depth of the troposphere and within the atmospheric boundary layer (ABL).

Assuming the requirement for tropospheric instability (or criticality; Houston and Niyogi 2007) is met, the potential for TI will remain sensitive to ABL characteristics and processes. Observational and modeling studies have examined the sensitivity of TI to ABL water vapour availability and depth (e.g., Mueller et al. 1993; Crook 1996; Weckwerth et al. 1996; Craven et al. 2002; McCaul and Cohen 2002), lift and depth of lift along convergence lines (e.g., Wilson et al. 1992; Ziegler and Rasmussen 1998; Crook and Klemp 2000; Wilson and Roberts 2006), and vertical wind shear in / near the ABL (e.g., Rotunno et al. 1988; Wilson et al. 1992; Mueller et al. 1993; Crook 1996; Crook and Klemp 2000; Markowski et al. 2006). Additional emphasis has been placed on the influence of mesoscale circulations on TI (e.g., Wilson et al. 1992; Crook 1996) and interactions between circulations and boundaries (e.g., Weckwerth and Wakimoto 1992; Sills et al. 2004; Buban et al. 2007). An excellent review of convective initiation mechanisms is given by Weckwerth and Parsons (2006). Many of the processes described in the above studies vary on small spatial and temporal scales that are not readily resolved using operational observation networks.

Sometimes, forecaster assessment of the pre-storm environment identifies the potential for severe thunderstorms but delineation of a TI threat area remains challenging. In Canada, contributing factors to this can include surface observations with coarse spatial and temporal resolution, widely-spaced and infrequent rawinsonde and AMDAR observations (in Canada most AMDAR data lack humidity observations), no operational profiling network, and limited Doppler radar coverage capable of detecting finelines via clear-air echoes. In short, access to observational data required to forecast TI can be limited. Canadian forecasters often rely on appropriate conceptual models and intelligent use of numerical weather prediction (NWP) output to fill in data-void areas. This paper describes work underway at Environment Canada's (EC's) Hydrometeorology and Arctic Lab (HAL) to generate real-time forecasts of TI using Canadian operational NWP output.

The following sections discuss convective forecasts in Canada, development of experimental model fields related to TI, formulation of TI forecasts, and subjective / objective verification of the forecasts. We conclude with a summary discussion and plans for future work.

2. CANADIAN NWP FORECASTS OF CONVECTIVE PRECIPITATION

The primary source of NWP data for Canadian forecasters is the Canadian Meteorological Centre's Global Environmental Multiscale (GEM) model (Côté et al. 1998). The GEM model is run regionally at 15-km horizontal grid spacing

[*] *Corresponding author address:* Neil M. Taylor, Hydrometeorology and Arctic Lab, Environment Canada, #200 4999 – 98th Ave. Edmonton, AB, T6B 2X3; e-mail: Neil.Taylor@ec.gc.ca

[**] Thunderstorm initiation (TI) is used in place of convective initiation (CI) to avoid subtleties that may be associated with the definition of CI. Results in this study are compared to lightning so that TI is less ambiguous.

(GEM15) over most of North America utilizing 58 levels in the vertical. Operational GEM15 forecasts are produced every 6 h out to T+48 h with output provided to forecasters at temporal resolutions of 3 h or 6 h depending on data requirements. Forecasts of convective precipitation (parameterized using the Kain-Fritsch scheme; Kain and Fritsch 1990) and other parameters are available operationally via forecaster workstations or web-based tools.

Public forecasts are generated at specific geographic locations via forecaster interaction with a meteorological database (SCRIBE) utilizing a variety of direct and post-processed GEM15 output. Using the SCRIBE graphical interface, meteorologists can edit forecast weather elements using the automatically-generated element in the database (or previous actual forecast) as a starting point. This allows the forecaster to consider trends in observational data and adjust the automatically-generated weather element concept accordingly. Forecasts of thunderstorms are produced in SCRIBE using a simple combination of GEM15 forecast Showalter Index (SI) and Updateable Model Output Statistics (UMOS; Wilson and Vallée 2002) forecasts of probability of precipitation (PoP). If a forecast of SI < 0 and PoP ≥ 30% occurs in a forecast region then a thunderstorm forecast element is generated. Environment Canada is exploring migration to an area-based forecast production system. Such a system will require spatial 'first-guess' forecast fields for TI and other weather elements in addition to existing point-based guidance.

Objective NWP forecasts of the timing, location, and intensity of warm-season convective precipitation tend to exhibit limited skill (e.g., Weckwerth et al. 2004; Wilson and Roberts 2006; Weisman et al. 2008). This is especially true for models with horizontal grid-spacing on the order of 10 km or greater (Weisman et al. 2008). While the operational GEM15 is not expected to pinpoint the timing and location of convective storms, characteristics of the general pre-storm environment can often be reasonably simulated. This can even include model representation of convergence boundaries (e.g., the dryline in Alberta). Following a rigorous assessment of observational data, forecasters may consult and modify NWP output to fit observations and actual expected trends. In this way, the forecaster can leverage the strengths of NWP, utilize aspects that are consistent with observations, adjust output for spatial or temporal deficiencies, and generate or modify a forecast. Utilizing the strengths of both the human forecaster and computers in the forecast production system allows for generation of appropriate convective forecasts (including watch / warning decisions) without depending exclusively on objective NWP solutions (see Sills (2009) for an excellent discussion on the role of the human forecaster).

3. EXPERIMENTAL GEM15 MODEL FIELDS

A variety of forecast parameters and indices have been developed to assist the meteorologist in assessing the potential for severe storms. These include characterizations of stability (e.g., MLCAPE, MUCAPE, 0-3 km CAPE, Normalized CAPE), deep vertical wind shear (e.g., 0-6 km bulk shear, Effective bulk shear), low-level shear and helicity (0-1 km bulk shear, 0-1 km SRH, Effective SRH) and composite indices (e.g., Energy-Helicity Index, Vorticity Generation Parameter, Supercell Composites, Significant Tornado Composites). Conversely, few forecast parameters are available that address the issue of TI specifically. To explore the use of GEM15 data for forecasting CI, a number of post-processed experimental NWP fields have been developed representing model-simulated ABL characteristics. During summer, fields are produced in real-time using the latest GEM15 REG model run and viewed via an internal web page designed for that purpose. Output is produced for each forecast hour from T+1 h to T+48 h over various domains in Canada. These experimental fields may be grouped in terms of water vapour availability and depth, convergence and convergence depth, and ABL vertical wind shear. Selected experimental fields are introduced below and illustrated for the T+6 h forecast valid 1800 UTC 30 July 2010 over southern Alberta (AB) and Saskatchewan (SK).

3.1 *Water Vapour Availability and Depth*

Under similar conditions of instability, lift, and surface water vapour mixing ratio (q_v), an environment with ABL water vapour mixed over a deep layer should be more conducive to surface-based TI than if ABL moisture is confined to a shallow near-surface layer (Mueller et al. 1993; McCaul and Cohen 2002). On the Canadian Prairies, especially in Alberta (Canada's equivalent to the U.S. High Plains), ABL water vapour can sometimes be a limiting factor for TI. Operationally, forecasters routinely characterize surface water vapour in terms of dewpoint (T_d). Anecdotal experience by the authors and forecasters is that GEM15 forecast surface T_d values are often higher than observed. Moreover,

model profiles suggest that surface values of T_d are often not representative of a mixed moist layer in the ABL (i.e., skin-layer moisture).

Two experimental fields discussed here characterize GEM15 representation of water vapour depth. The first is a depiction of the T_d associated with a mixed layer of moisture over the lowest 50 hPa of the model (AVTD; Fig 1). A 50 hPa layer is commonly used on the prairies for lifting parcels on tephigrams to assess static stability.

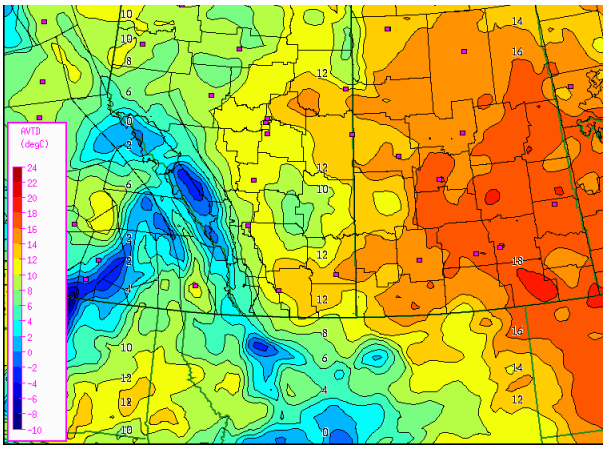

Fig. 1. AVTD for 1800 UTC 30 July 2010.

A second representation of GEM15 ABL moisture depth is a mixed moist layer depth (MMLD) parameter displayed along with corresponding mixing ratio (Fig. 2). The top of the mixed layer is determined by identifying where GEM15 mixing ratio decreases rapidly at successive levels above the surface.

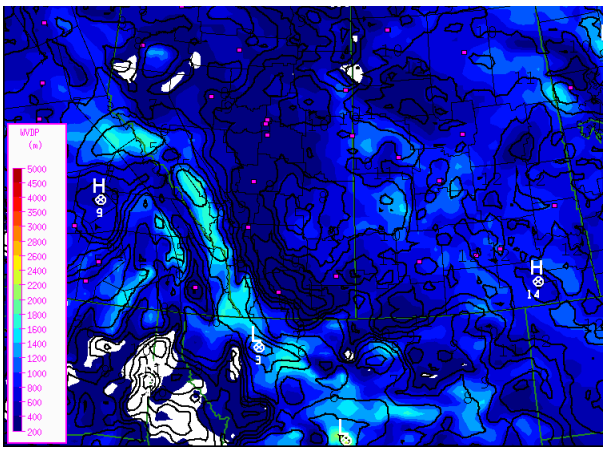

Fig. 2. MMLD and associated q_v for 1800 UTC 30 July 2010.

The examples shown here illustrate a region with high AVTD (16-18 °C) over southern SK with decreasing values (10-12 °C) through the eastern portions of AB. Over the AB foothills, much lower AVTD values predominate with a developing dryline indicated by the gradient in AVTD separating the two areas. In Fig. 2, MMLD (q_v) values are near 1000 m (10 g kg^{-1}) or higher in SK and only 600 m (9 g kg^{-1}) or less over eastern AB. Over the foothills the MMLD is much deeper (~1600 m) but with q_v values of only ~ 5 g kg^{-1} or less. In this case the contrasts in ABL moisture across the developing dryline are clearly illustrated.

These fields can help forecasters identify areas where GEM15 forecast ABL moisture is confined to a thin layer vs. well-mixed over a deeper layer. They can also be useful for conceptualizing GEM15 representation of synoptic or mesoscale features, e.g., contrasts in moisture availability and depth across the dryline or surface fronts.

3.2 *Convergence and Convergence Depth*

Low-level convergence zones are recognized as favoured regions for TI (e.g., Wilson and Schreiber 1986; Wilson et al. 1992; Weckwerth and Parsons 2006). The effects of lift and convergence above the surface have also been found to influence TI along drylines (Ziegler and Rasmussen 1998) and with respect to divergence above the convergent layer (Wilson et al. 1992; Crook 1996). In general, areas with low-level convergence over a deep layer should be more favoured for TI than an area with similar (or weaker) convergence over a shallow layer. Experimental NWP fields have been developed to identify forecast areas of low-level convergence and depth of the convergence layer. The top of the convergence layer is identified as the level above the surface where the sign of divergence changes from negative (i.e., convergence) to positive (i.e., divergence). Surface divergence (SDIV) and wind are illustrated in Fig. 3 while the depth of the convergence layer (COND) is shown in Fig. 4.

The SDIV field (Fig. 3) highlights an organized area of convergence along the AB Foothills with less-organized convergence areas over SK. The small areas of surface divergence (and accompanying nearby convergence areas) in SK indicate model-simulated convective cells within a weak convergence zone. In Fig. 4, COND clearly highlights a region of deep convergence (COND > 1500 m) associated with the model-simulated dryline over the AB Foothills. Smaller areas of similar COND values are indicated over SK but COND is generally ~ 1250 m or less.

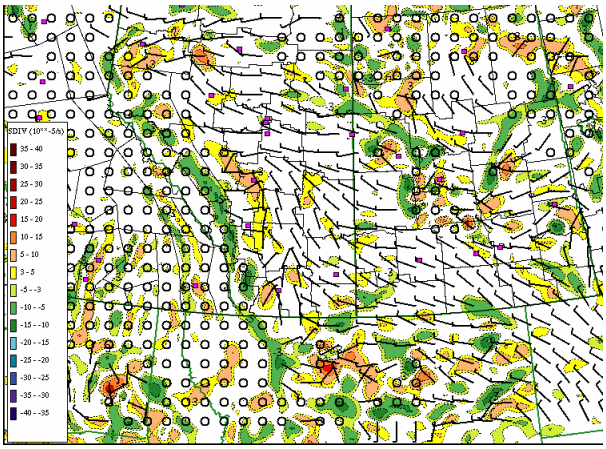

Fig. 3. SDIV and surface wind for 1800 UTC 30 July 2010.

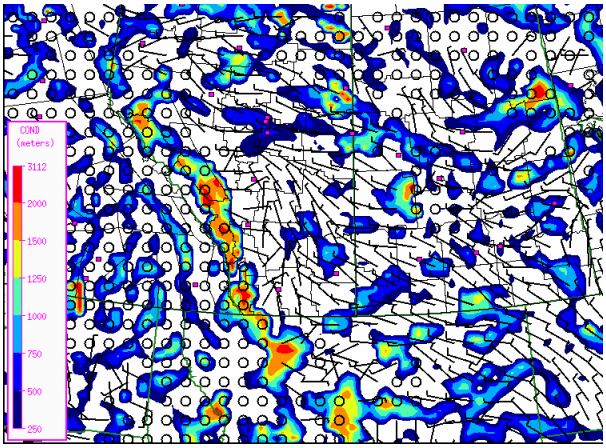

Fig. 4. COND for 1800 UTC 30 July 2010.

An additional consideration for TI with respect to convergence depth is the height of the level of free convection (LFC). To examine the depth of the convergence layer relative to the LFC, a field was developed showing the ratio of COND to the 50 hPa mean layer (ML) LFC height (RCLFC; Fig. 5).

Areas where the depth of convergence extends to, or higher then, the height of the MLLFC will have RCLFC ≥ 1. Areas with higher RCLFC should be more favoured for TI. Values of RCLFC in Fig. 5 are similar over the AB Foothills and SK (generally > 0.3 with some areas > 0.6 and smaller areas > 1.0). Note the forecast MLLFC height over SK was ≤ 1500 m compared to values ≥ 2000 m over AB (not shown). Based on SDIV, COND, and RCLFC considerations alone, favorable areas for TI would include the AB foothills and along a generally NW to SE axis over southern SK.

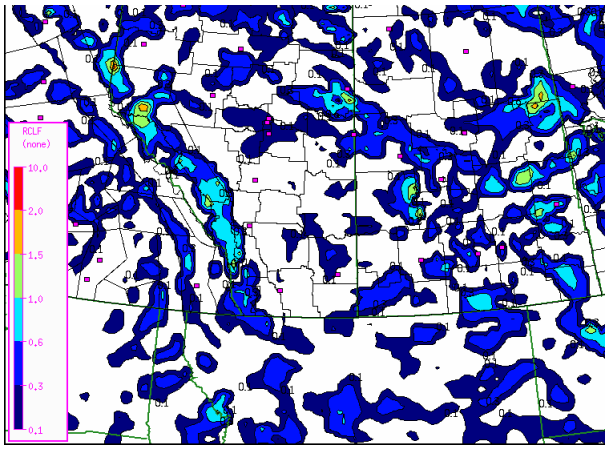

Fig. 5. RCLFC for 1800 UTC 30 July 2010.

3.3 *ABL / Near-ABL Wind Shear and Buoyancy*

Low-level convergence boundaries are recognized as a mechanism to promote vertically-oriented updrafts that penetrate well into the ABL (Rotunno et al. 1998). When vertical wind shear is present across a convergence boundary, parcel updrafts can be tilted in the down shear direction. If this occurs below the lifting condensation level (LCL), the increased residence time of the parcel in the ABL allows for increased dry-air entrainment that may reduce positive buoyancy and delay, or inhibit, TI (Wilson et al. 1992; Mueller et al. 1993; Crook 1996; Markowski et al. 2006).

Bulk shear from the most unstable lifted parcel level (MULPL) to the MULFC ($SHR_{0\text{-}LFC}$) is displayed in a field along with wind barbs representing the shear vector (Fig. 6). TI in areas with strong vertical shear oriented across a convergence boundary may be delayed. It is recognized that while initially delaying TI and storm organization, strong vertical shear at low levels is favorable for the development of severe storms, e.g., tornadic supercells. As such, $SHR_{0\text{-}LFC}$ bulk shear may also be useful as a tool for diagnosing severe storm potential. In Fig. 6, much of southern SK is characterized by weak, or no, $SHR_{0\text{-}LFC}$ while over the AB Foothills values of $SHR_{0\text{-}LFC}$ are ≥ 20 kt.

The effect of vertical shear delaying convection may also occur above the LFC. If vertical shear is strong in some layer above the LFC it may be expected that organization of an incipient storm could be delayed. Another field developed was the bulk shear in the lowest 2 km above the MULFC (SHR_{LFC+2}) displayed along with CAPE from the MULPL to 3 km above the MULPL (Fig 7). For surface-based convection, this reduces to the familiar 0-3 km CAPE. The region over southern SK is shown in Fig. 7 to be characterized by higher

MULPL-3 km CAPE than most of AB but with SHR_{LFC+2} values similar in both regions.

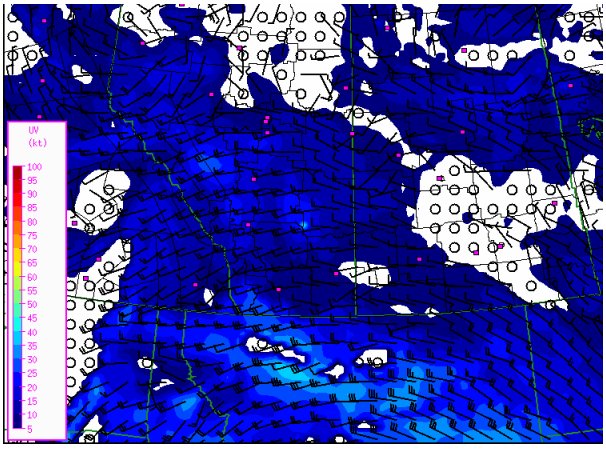

Fig. 6. $SHR_{0\text{-}LFC}$ for 1800 UTC 30 July 2010.

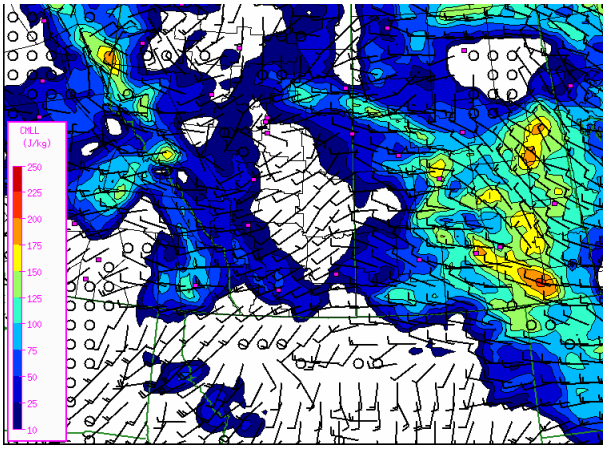

Fig 7. SHR_{LFC+2} and MULPL-3 km CAPE for 1800 UTC 30 July 2010.

Consideration of Figs. 1-7 suggests two regions might be targeted by the forecaster for TI (actual potential for TI will be contingent on further consideration of stability / convective inhibition (CIN) and comparison with observed data). A developing dryline over the AB Foothills is represented by strong contrasts in ABL moisture and convergence over a deep layer (nearly as high as the MLLFC). Given only moderate values of AVTD, and fairly shallow MMLD, water vapour availability may be a limiting factor for TI. Wind shear within and near the ABL may act to delay, or possibly inhibit, TI altogether. Over southern SK, general conditions of high AVTD and MMLD values, local areas of deep convergence relative to the MLLFC, and generally weak ABL / near ABL vertical wind shear should readily favor TI.

For the purposes of illustration, forecast AVTD for 1800 UTC 30 July 2010 is shown with observed CG lightning from 1800:00 to 1859:59 in Fig. 8. The few flashes over the AB Foothills indicate TI along the forecast position of the dryline (this storm would later evolve into a tornadic supercell) and a weak storm crossing the mountains to the SW. Over SK, most of the lightning appears to be associated with remnant elevated storms while the southernmost lightning flashes are associated with new surface-based storms (more surface-based storms would later develop in this region). The lightning in west-central AB appears to be associated with surface-based storms in the vicinity of a low pressure circulation.

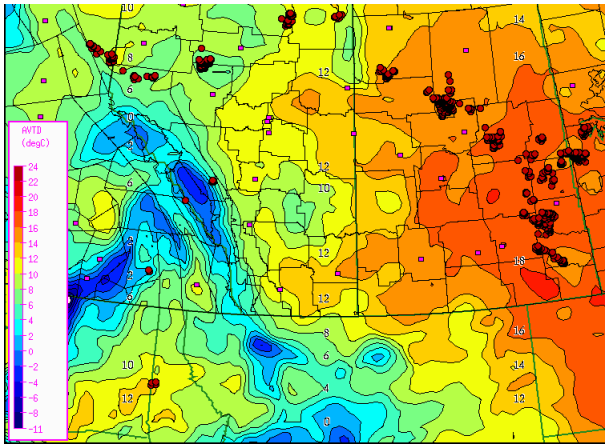

Fig. 8. AVTD and observed lightning (red circles) for 1800 UTC 30 July 2010.

3.4 *Comparison with Observed Lightning*

In sections 3.1-3.3, one case was used to illustrate how selected experimental fields could be interpreted in an operational setting. These and other fields were tested in real time on the Research Support Desk (RSD; see Sills and Taylor 2008) in Prairie and Arctic Storm Prediction Centre (PASPC)-Edmonton operations during the summer of 2007. They were found at times to be useful visualization tools for model behaviour when defining threat areas for TI. For a more general comparison with observed convection, data from the Canadian Lightning Detection Network (CLDN) was used. CLDN data were chosen as the best objective indicator of thunderstorms given limited radar coverage and manned observations across the Canadian prairie domain.

A number of GEM15 parameters (1200 UTC run only) were compared with observed cloud-to-ground flashes between 1500 UTC and 0300 UTC each day from 1 May to 30 September 2007 (517000+ flashes). This time period was selected

to encompass the typical daytime TI period and minimize the frequency of flashes from elevated nocturnal storms. Model fields were assumed valid for the duration of the forecast hour so that lightning data from the field valid time (T) to T+59:59 minutes were used for each forecast hour (T+3 to T+15)*. Since we have not attempted to separate flashes from surface-based vs. elevated storms, in some cases near-surface parameter values may not be representative of the actual storm environment. Selected results are discussed below with basic descriptions of each distribution in Table 1.

A histogram of AVTD values is shown in Fig. 9. The mean (median) value is 12.3 °C (12.7 °C) with the bulk of the distribution (~ 90 %) falling into the range of 5-18 °C and peaking at 16-17 °C. When examining MMLD, the associated mixing ratio has to be considered to separate environments with a deep and dry ABL from a shallow and moist ABL (e.g., across a dryline). A plot of MMLD and associated mixing ratio is shown in Fig. 10. MMLD values ranged from near zero (40 m) to a maximum near 4200 m AGL with associated mixing ratios ranging from 3-20 g kg^{-1}. Mean (Median) values of MMLD are 1023 m (965 m) with 90% of the distribution in the range 400-2000 m and the peak of the distribution in the 800-1000 m range (not shown). The majority of flashes (96%) were associated with a MMLD less than 2000 m and q_v in the range 4 to 16 g kg^{-1} (not shown). Very few flashes were associated with MMLD more than 2500 m and those that were had q_v less than ~ 8 g kg^{-1}.

A plot of COND (including flashes associated with low-level divergence) and SDIV is shown in Fig 11. The majority of observed flashes (60%; not shown) were associated with convergence over some depth but a significant number of flashes (40%; not shown) were associated with forecast low-level divergence. This is likely due to incorrect model placement of convergence areas relative to actual storms and to some fraction of storms being elevated and not necessarily tied to surface-based convergence areas.

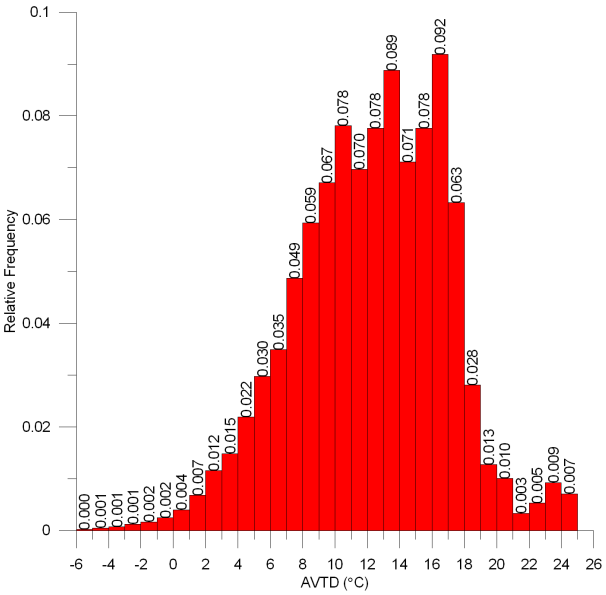

Fig. 9. Histogram of AVTD.

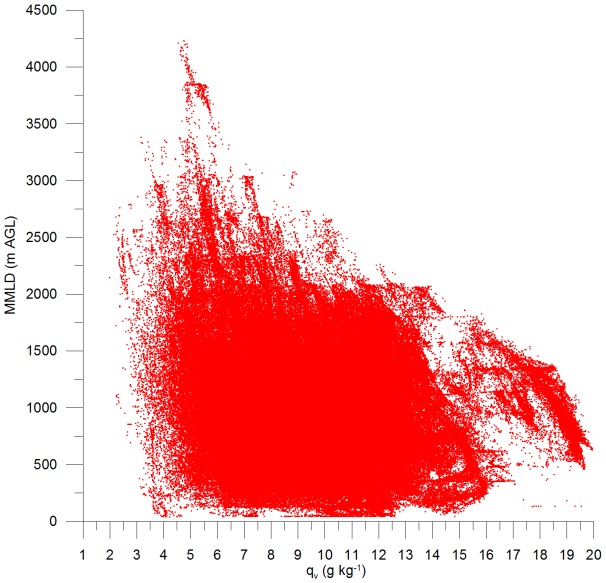

Fig. 10. plot of MMLD and associated mixing ratio.

* Strictly speaking, only the initial flash for a given storm should be used for comparison. However, objective identification of 'first flash' lightning over an entire season is difficult. For these preliminary results, we have approximated TI by using multiple flashes to facilitate the analysis and maximize the sample size of flashes used for comparison.

Table 1. Experimental GEM15 model parameters and basic description of the distribution of values when compared to observed lightning in 2007.

Parameter	Maximum	Minimum	1st Quartile	3rd Quartile	Mean	Median	Std. Dev.
AVTD (°C)	24.8	-9.7	9.2	15.8	12.3	12.7	4.7
MMLD (m)	4229	40	662	1319	1023	965	496
SDIV (x10^{-5} s^{-1})	40.2	-41.8	-4.3	1.8	-1.4	-1.0	6.0
COND (m)	4129	0	487	1220	899	830	549
RCLFC	10.0	0	0	0.3	0.2	0.1	0.2
$SHR_{0\text{-}LFC}$	70.8	0	4.1	25.3	16.4	14.8	13.7
SHR_{LFC+2}	54.7	0	2.3	13.6	8.9	7.9	7.6

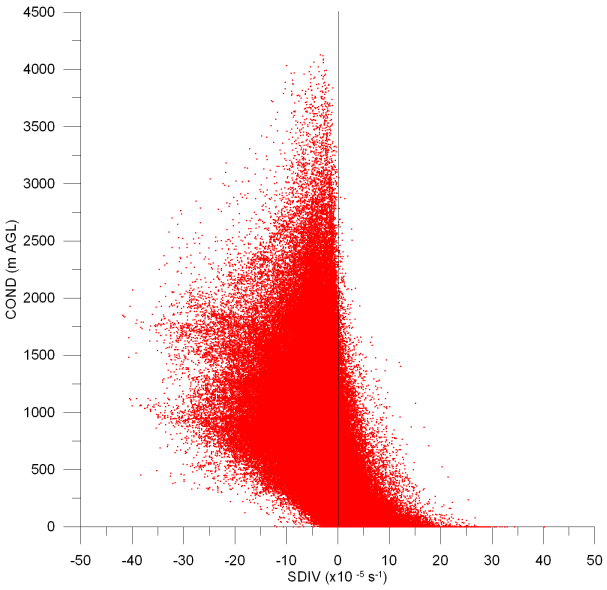

Fig. 11. Plot of COND (including flashes associated with low-level divergence) and SDIV. The vertical line separates convergence (SDIV < 0) and divergence (SDIV > 0).

The depth of the layer for divergence-associated flashes is generally confined to the lowest ~1500 m. For flashes associated with low-level convergence, COND values occur as high as 4000 m. Approximately 90% of COND values are in the 200-2000 m range with the peak in the distribution is in the 600-800 m range (not shown). There appears also to be a relationship between COND and the magnitude of convergence as stronger convergence (greater than ~ 15 x10^{-5} s^{-1}) tends to be associated with depths in the range 500 m to 2500 m.

The distributions of $SHR_{0\text{-}LFC}$ and SHR_{LFC+2} values are shown in Fig. 12 and Fig. 13, respectively. Both shear parameters tend to have higher relative frequencies for low values of bulk shear. Over 60% of lightning flashes are associated with $SHR_{0\text{-}LFC}$ less than 20 kt and ~ 60% of flashes are associated with SHR_{LFC+2} less than 10 kt.

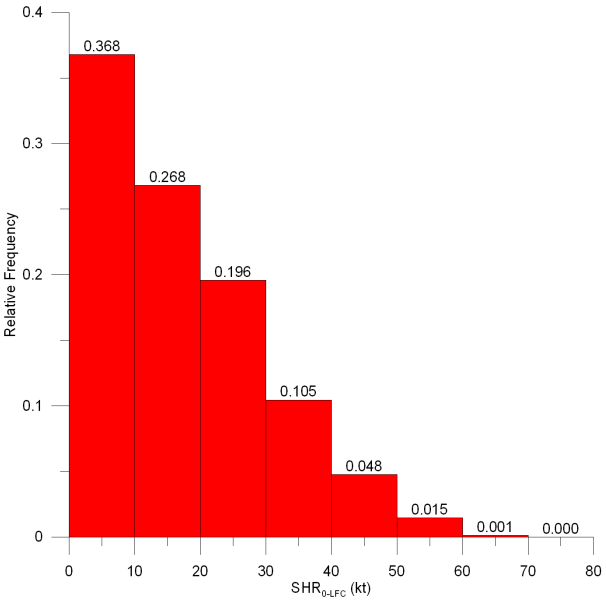

Fig. 12. Histogram of $SHR_{0\text{-}LFC}$.

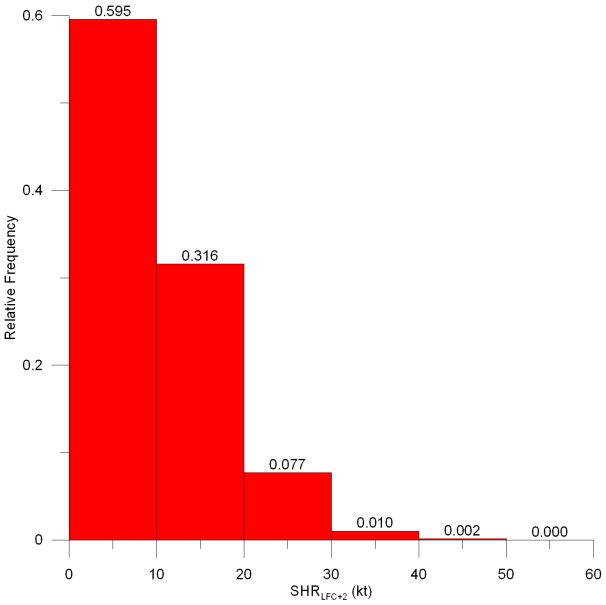

Fig. 13. Histogram of SHR_{LFC+2}.

The non-random distributions of experimental parameter values suggest that some relationships exist between observed lightning and the

experimental GEM15 model fields examined here. The distributions are well-defined with peaks in ranges that seem appropriate given our understanding of TI processes and conceptual models. Forecast AVTD values associated with observed lightning are distributed broadly with most values in the 5-18 °C range. MMLD values are more narrowly distributed with the bulk of values (80%) in the range 400-1600 m range. There is a tendency for lightning to be associated more often with convergence than divergence and 60% of COND values are in the range of 400-1400 m. Observed lightning also tends to be associated with low forecast values of $SHR_{0\text{-}LFC}$ and SHR_{LFC+2}. These results suggest there may be some utility in using these experimental parameters in forecasts of thunderstorms and TI.

4. DEVELOPMENT AND VERIFICATION OF EXPERIMENTAL TI FORECASTS

There is building momentum within Environment Canada to move towards an area-based forecast production system. Such a system would benefit from the addition of spatial 'first-guess' fields of TI, and other weather elements, in addition to UMOS-generated point forecasts. Using results described in section 3.4, a number of preliminary thresholds for selected forecast parameters were subjectively defined and used to develop a set of TI forecasts. These were composite forecasts using selected experimental parameters in combination with forecasts of static stability and CIN. We used an ingredients-based approach to develop the forecasts considering basic requirements for convection (sufficient moisture, instability, and lift) and ABL characteristics that influence TI.

Five objective forecasts were initially developed and evaluated to account for the effects of ABL water vapour availability / depth (MDEP), convergence / convergence depth (CNVD), low-level shear and CAPE (LLWC), and CIN (CVIN). Included was an overall objective TI forecast (OBTI) combining various elements of the above. Forecasts were produced and made available to PASPC operations via an internal web page during the 2009 convective season. The forecasts were evaluated in real time against observed lightning on the RSD. For comparison, other objective forecasts related to convective storms were also evaluated. These included GEM15 convective precipitation rates from the Kain-Fritsch scheme (KF), a spatial representation of the UMOS-generated SCRIBE thunderstorm forecast (pseudo-SCRIBE or P_SCRIBE), a representation of the integrated vertical velocity below the ML equilibrium level (WSUM)[*], and a modified version of the Cloud Physics Thunder Parameter (CPTP) developed and used at the Storm Prediction Center in Norman, OK. These objective forecasts are not TI forecasts per se (i.e., KF and WSUM do not necessarily imply lightning and CPTP is a conditional forecast assuming TI) but were the best available objective forecasts of convection available for comparison. The forecasts verified are listed in Table 2[**] along with thresholds of their component parameters.

Each objective TI forecast was developed with multiple thresholds to highlight regions that, based on GEM15 model characteristics, would have an increased likelihood for TI. Each threshold was assigned a color (blue, green, yellow, red) so that warmer colors indicated more stringent parameter value thresholds and an increased likelihood for TI. The forecasts were subjectively verified against observed lightning in 3 phases each lasting 3 weeks. At the end of each phase, parameter thresholds were adjusted to address observed deficiencies in the forecasts. The formulations at the end of the third phase were considered final.

4.1 *Subjective Verification of TI Forecasts*

Real-time subjective verification of TI forecasts was conducted to assess the utility of the forecasts in an operational setting. A subjective verification was utilized to observe not only if forecasts were successful or failed, but why (e.g., poor forecast formulation vs. model performance). It was recognized that model placement of convergence lines and other features would not be consistently accurate. It was also recognized that the forecaster can adjust for timing and location of model features using observational data and still extract useful information from model forecasts even if placement of a feature is not correct.

[*] During subjective verification, scores for WSUM were similar to (though slightly poorer than) KF scores so verification of WSUM was not included in the third phase.

[**] MDEP and CNVD in Table 2 do not include a condition for instability and were therefore not verified as TI forecasts. They were evaluated as combined representations of model forecast moisture and convergence depth only.

Table 2. Objectively verified TI and existing convection forecasts, composite parameters used, and thresholds for blue (B), green (G), yellow (Y), and red (R) forecasts. Positive (negative) thresholds are exceeded when the model value is ≥ (≤) the listed value. Units are given in square brackets.

TI Forecast	Parameters	Thresholds (B, G, Y, R)
Moisture Availability and Depth (MDEP)	AVTD MMLD	4, 6, 8, 10 [°C] 400, 600, 800, 1000 [m]
Convergence and Convergence Depth (CNVD)	SDIV COND RCLFC	-2.5, -5.0, -7.5, -10 [x10^{-5} s^{-1}] --, 500, 750, 1000 [m] --, --, 0.25, 0.5
Low-Level Shear and CAPE (LLWC)	MUCAPE (MLCAPE) MULPL-3km CAPE $SHR_{0\text{-}LFC}$ SHR_{LFC+2}	100 (50), 200 (100), 300 (200), 400 (400) [J kg^{-1}] 25, 50, 75, 100 [J kg^{-1}] 40, 30, 20, 10 [kt] 25, 15, 10, 5 [kt]
Convective Inhibition (CVIN)	MUCAPE (MLCAPE) MUCIN MULFC-MULCL	100 (50), 200 (100), 300 (200), 400 (400) [J kg^{-1}] -100, -50, -25, -10 [J kg^{-1}] 1000, 750, 500, 250 [m]
Objective TI Forecast (OBTI)	MUCAPE (MLCAPE) MULPL-3km CAPE MUCIN MULFC-MULCL SDIV COND RCLFC	100 (50), 200 (100), 300 (200), 400 (400) [J kg^{-1}] --, 50, 75, 100 [J kg^{-1}] -100, -50, -25, -10 [J kg^{-1}] 1000, 750, 500, 250 [m] --,--, -2.5, -5.0 [x10^{-5} s^{-1}] --, 500, 750, 1000 [m] --, --, --, 0.5
Kain-Fritsch Rain Rate (KF)	Direct from model	0.1 x10^{-7} [m s^{-1}]
Pseudo SCRIBE (P_SCRIBE)	SI Total model precipitation	Negative 0.2 [mm]
Modified CPTP (CPTP)	Bright et al. (2005)	25
Total GEM15 Vertical Velocity Below the ML Equilibrium Level (WSUM)	--	--

In our method, two raw scores were defined to consider all four quadrants of a standard 2x2 contingency table (Fig. 14). A HIT-MISS (HM) score was assigned according to:

(a) observed lightning within a positive forecast area (a hit), and
(b) observed lightning outside a positive forecast area (a miss).

		Observed	
		YES	NO
Forecast	YES	*HIT*	*FALSE ALARM*
	NO	*MISS*	*CORRECT NULL*

Fig. 14. A standard 2x2 contingency table used for verification of yes/no forecasts.

HM scores were improved (degraded) for more observed lightning inside (outside) positive forecast areas. Based on the areal extent of lightning and forecast areas, an overall score out of 5 was assigned. A False Alarm-Correct Null (FA) score was assigned according to:

(a) positive forecast areas with no lightning observed (a false alarm), and
(b) null forecast areas where no lightning was observed (a correct null forecast).

As with the HM score, the FA score was assigned a value out of 5 based on the areal extent of positive and null forecast areas over the prairies*. If no lightning was observed, no HM score could be awarded but a FA score was still assigned. An overall combined score was defined as the sum of the HM and FA scores. A perfect forecast would therefore have a combined score of 10.

A graph of HM and FA scores for the third verification phase is given in Fig. 15. For all the experimental TI forecasts, HM (FA) scores were degraded (improved) as thresholds were increased in the blue (B), green (G), yellow (Y) and red (R) forecasts. The result is that the blue

* Note that for FA scores, higher values indicate a better forecast.

scores for the experimental TI forecasts were consistently highest. Of the KF, P_SCRIBE and CPTP scores, KF and CPTP have similar performance to the blue TI forecasts but the P_SCRIBE forecast, while having a good FA score, suffers from a poor HM score. The combined scores are shown in Fig. 16 indicating LLWC_B as the best performing forecast overall. A ranking of the forecasts based on combined score is given in Table 3.

The author's impressions when conducting the subjective verification was that a number of the forecasts served useful as guidance for forecasting TI. Often, areal coverage of a forecast was good but the HM scores were limited by incorrect model placement of some features, most notably consistent locations of convergence lines. When any of the convergence-related parameters were used in the TI forecasts (e.g., OBTI_Y) the HM scores for detection suffered. Even in cases where vertical forcing was incorrectly located, however, other information on the convective environment remained useful for identifying TI areas. The difference in performance by the top forecasts (LLWC_B, OBCI_B, CVIN_B) and KF was small whereas all the forecasts performed significantly better than the P_SCRIBE forecast. This suggests that even a simple replacement of KF over the existing UMOS-derived SCRIBE point forecast may prove beneficial for first-guess forecasts in an area-based forecast system.

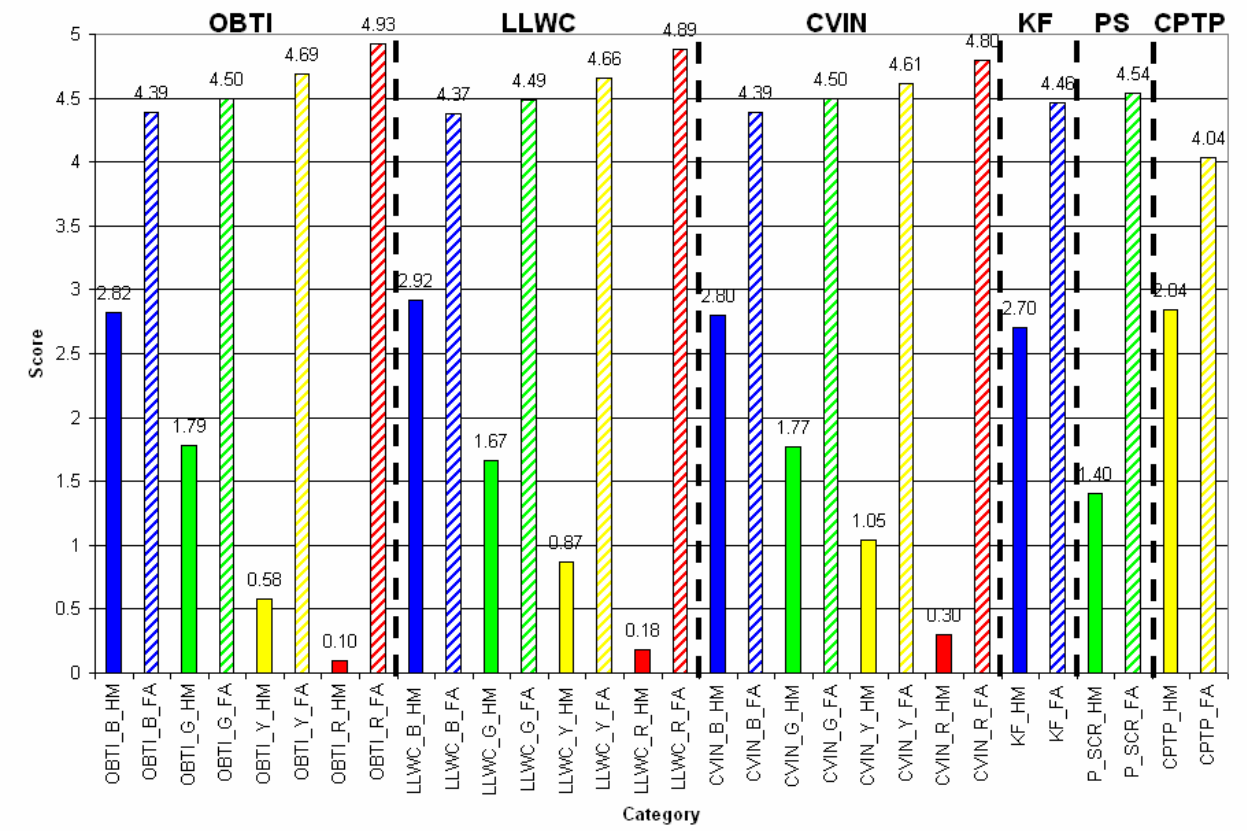

Fig. 15. Raw HM and FA scores from the third phase of subjective verification of experimental TI and existing objective convection forecasts.

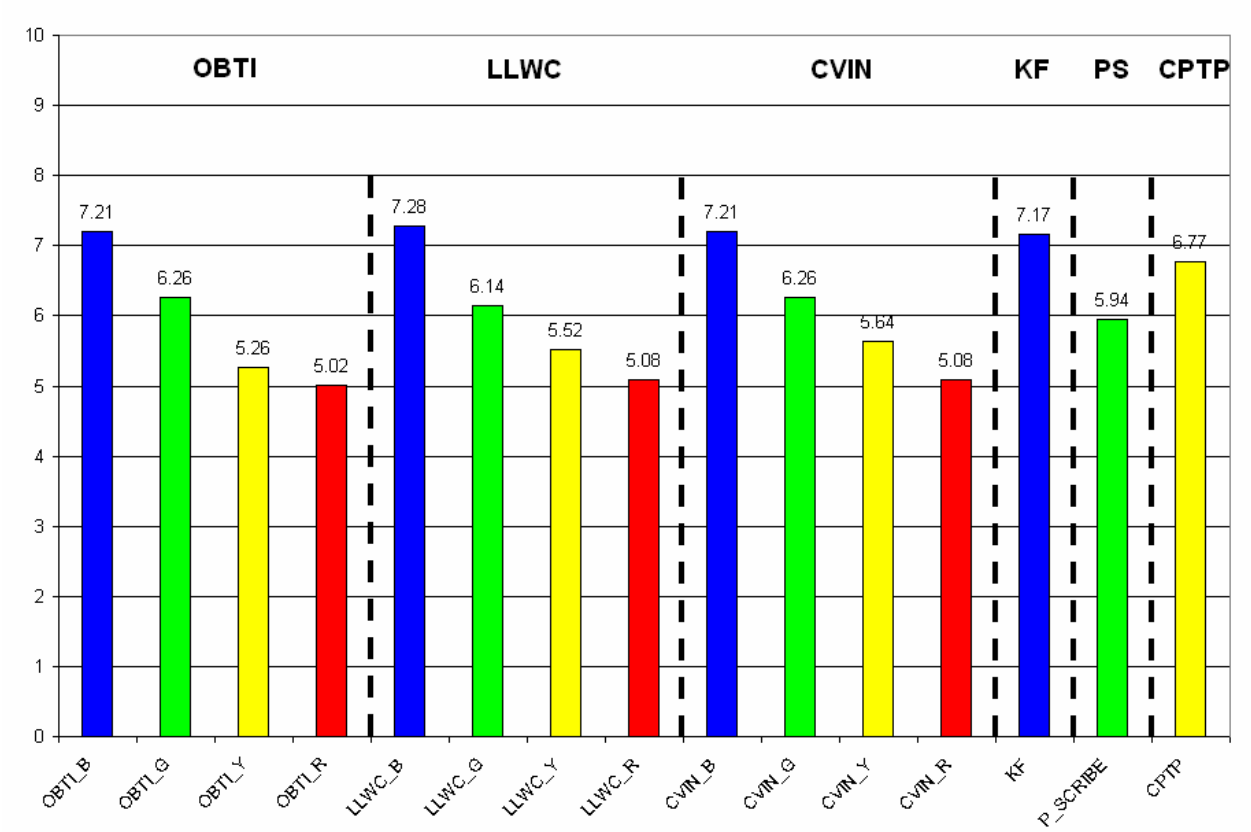

Fig. 16. Combined scores from the third phase of subjective verification of experimental TI and existing objective convection forecasts.

Table 3. Ranking of "blue" experimental TI and existing objective convection forecasts by combined score from subjective verification.

TI Forecast	Total HM Score	Total FA Score	Combined Score	Rank
LLWC	2.92	4.37	7.28	1
OBTI	2.82	4.39	7.21	2
CVIN	2.80	4.39	7.21	3
KF	2.70	4.46	7.17	4
CPTP	2.84	4.04	6.77	5
P_SCRIBE	1.40	4.54	5.94	6

4.2 *Objective Verification of TI Forecasts*

The subjective verification process was useful for assessing the utility of the TI forecasts in an operational environment and allowed for an iterative approach to modify forecast formulations. The subjective results may, however, be influenced by the small sample size and subjective nature of the verification (e.g., bias of the evaluator). For a complete description of forecast verification over the convective season an objective methodology should also be employed. We considered all the 1200 UTC run GEM15 TI forecasts from 1 May to 30 September in 2009. Forecasts in the T+1 h to T+18 h timeframe and observed lightning from the same period were used to conduct the objective verification. An attempt was made to account for some error in model placement of significant features. For each forecast grid point (i, j), the value of the forecast was extended to be valid at 2x the GEM15 horizontal grid spacing (i.e., 15 km) in the x (u) and y (v) directions. This effectively created a 5x5 grid-point tile with the forecast value determined at the central grid point (Fig. 17).

Any lightning flash observed within a positive forecast 5x5 tile renders the forecast a hit. Similarly, lightning observed outside the 5x5 tile is a miss, and no observed lightning inside the tile results in a false alarm. A null forecast tile with no lightning observed within the tile is a correct null forecast. Once a 2x2 contingency table was

populated for each grid point over the prairie provinces domain, a number of standard verification metrics (see Appendix A) were calculated to assess the skill of the forecast. The resulting overall scores are given in Table 4*.

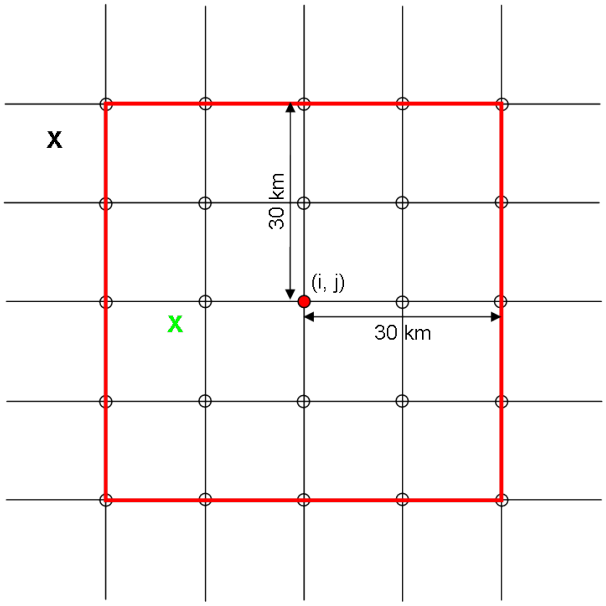

Fig. 17. Schematic showing the 5x5 grid-point tile used in the objective verification methodology. The red box represents the extent to which the forecast value at point (i, j) is valid for verification. The 5x5 grid-point tile covers a 60x60 km box in which lightning flashes would be counted. For a positive forecast, the green **x** would be a hit while the black **x** would be a miss.

The overall scores for the all the forecasts are not very impressive with generally low CSI, HK, and HSS scores. The scores in Table 4 are ranked by POD as it is felt that detection of potentially hazardous weather (lightning in this verification) is the most critical aspect of these forecasts. This is not to undervalue the importance of ensuring low false alarm scores, however, and it is noted that all the forecasts suffer from poor FAR scores. The existing objective convective forecasts WSUM, KF, and P_SCRIBE suffer from poor POD scores but have slightly improved FAR scores while CPTP performed better than OBTI_B. Even with the use of a 5x5 grid-point tile for verification, the size of the verification domain (the three prairie provinces) and relatively small areas associated with observed lightning may make it difficult to obtain high verification scores using this methodology.

* Note from Table 2 that the "blue" forecast of CVIN has the same criteria as for OBTI so CVIN_B is not included in further discussions of objective verification.

When the objective verification is conducted for each forecast hour individually, there is some improvement in scores for the daytime convective period. Hourly graphs of POD and FAR are shown in Figs. 18-19. There is a clear separation in POD scores (Fig. 18) for LLWC_B, CPTP, and OBTI_B compared to the other forecasts for the hours of T+4 h (1600 UTC) to T+14 h (0200 UTC). The existing objective convection forecast scores show little change throughout the verification period (T+1 to T+18 h). There is a small improvement in FAR scores (Fig. 19) over approximately the same period but the improvement is not as dramatic as for POD. This suggests that specific forecasts may be most appropriate for different time periods. In the morning and evening hours, there may be limited improvement in using the experimental TI forecasts over existing objective convection forecasts (e.g., KF). In the daytime TI period (~1600-0200 UTC) there may be more benefit in using an experimental TI forecasts (e.g., LLWC_B).

There is some uncertainty as to which verification metrics are most appropriate to consider when verifying a forecast as there are pros and cons to each. We have already stated our position that detection of hazardous weather is the most critical attribute of these forecasts. Ideally, a forecast can be designed that maximizes detection while minimizing false alarm forecasts. It is also important for forecasts of rare significant events, to consider correct null forecasts in overall scores. In our objective verification, the metrics that include correct null forecasts are HK, HSS, and ODDS. We are encouraged that the experimental CI forecasts (LLWC_B, OBTI_B) rank high in terms of POD and HK but note that they suffer in terms of HSS and ODDS. Given that HSS compares forecast accuracy to random chance, this may not be the best metric on which to base forecast decisions. The use of HK is attractive given that it assesses how well a forecast separates yes events from no events. For hourly forecasts, HK values of LLWC_B and CPTP peak at 0.59 and 0.56, respectively for the T+8 h (2000 UTC) forecast (not shown). These forecasts, along with OBTI_B score consistently higher than the other forecasts verified during the daytime TI period. From the subjective verification, LLWC_B and OBTI_B had the best overall scores suggesting that these forecasts, along with CPTP, may be the best forecasts to concentrate on improving for use in an area-based forecast production system. It should be noted that in both subjective and objective verification methods, the P_SCRIBE forecast was consistently among the

worst forecasts available. We acknowledge, however, that the P_SCRIBE forecast verified here does not benefit from the application of UMOS used in the generation of actual thunderstorm forecasts in SCRIBE.

Table 4. Overall objective verification scores for "blue" TI forecasts and existing objective convection forecasts ranked by POD.

Forecast	POD	FAR	CSI	ACC	POFD	HR_Y	HR_N	HK	HSS	ETS	BIAS	ODDS	POD/FAR
LLWC_B	0.682	0.885	0.109	0.835	0.161	0.115	0.989	0.522	0.154	0.083	5.935	11.224	0.771
CPTP	0.595	0.857	0.130	0.882	0.109	0.143	0.986	0.486	0.192	0.106	4.159	12.033	0.695
OBTI_B	0.583	0.880	0.110	0.860	0.131	0.120	0.986	0.452	0.157	0.085	4.873	9.254	0.662
WSUM	0.311	0.866	0.104	0.920	0.061	0.134	0.978	0.250	0.153	0.083	2.315	6.923	0.360
KF	0.306	0.798	0.138	0.943	0.037	0.202	0.978	0.269	0.215	0.121	1.519	11.467	0.384
P_SCRIBE	0.180	0.747	0.118	0.960	0.016	0.253	0.975	0.164	0.190	0.105	0.711	13.301	0.241

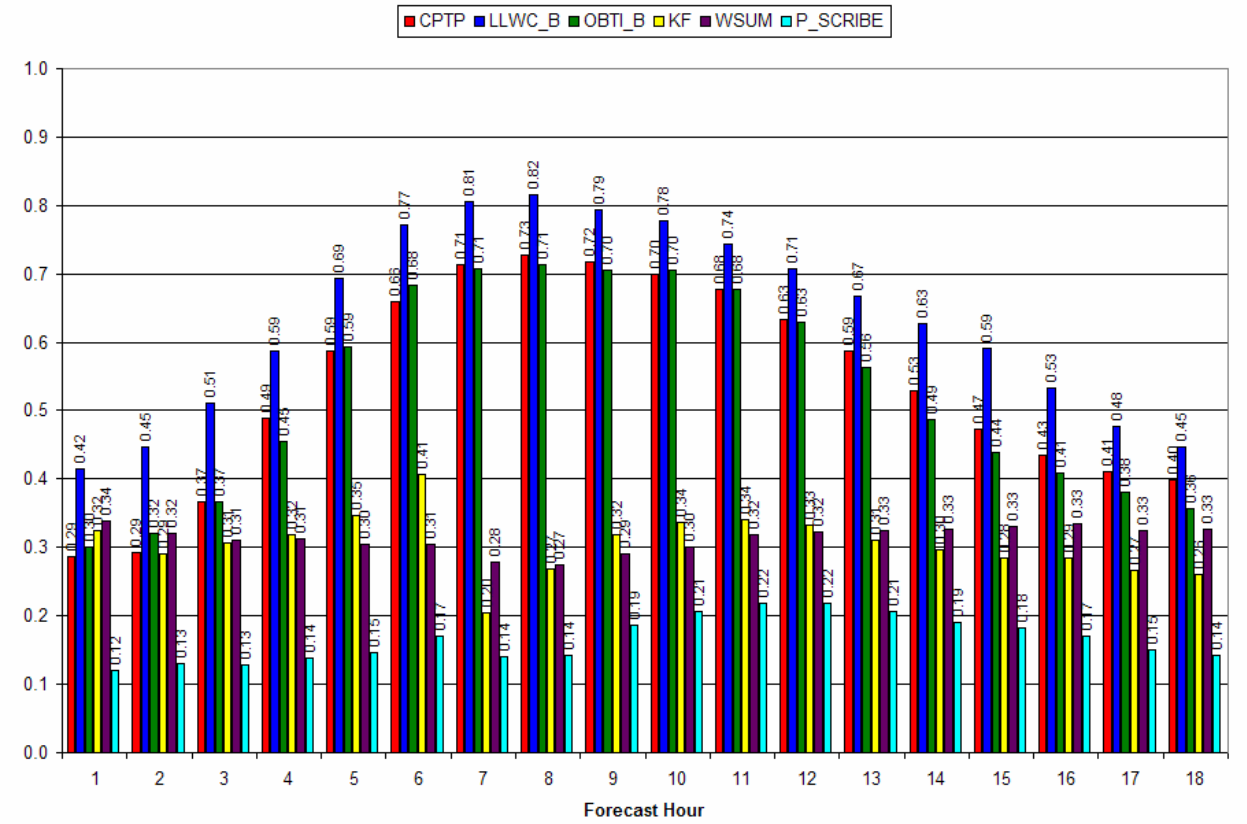

Fig. 18. Objective verification scores for POD from T+1 h (1300 UTC) to T+18 h (0600 UTC).

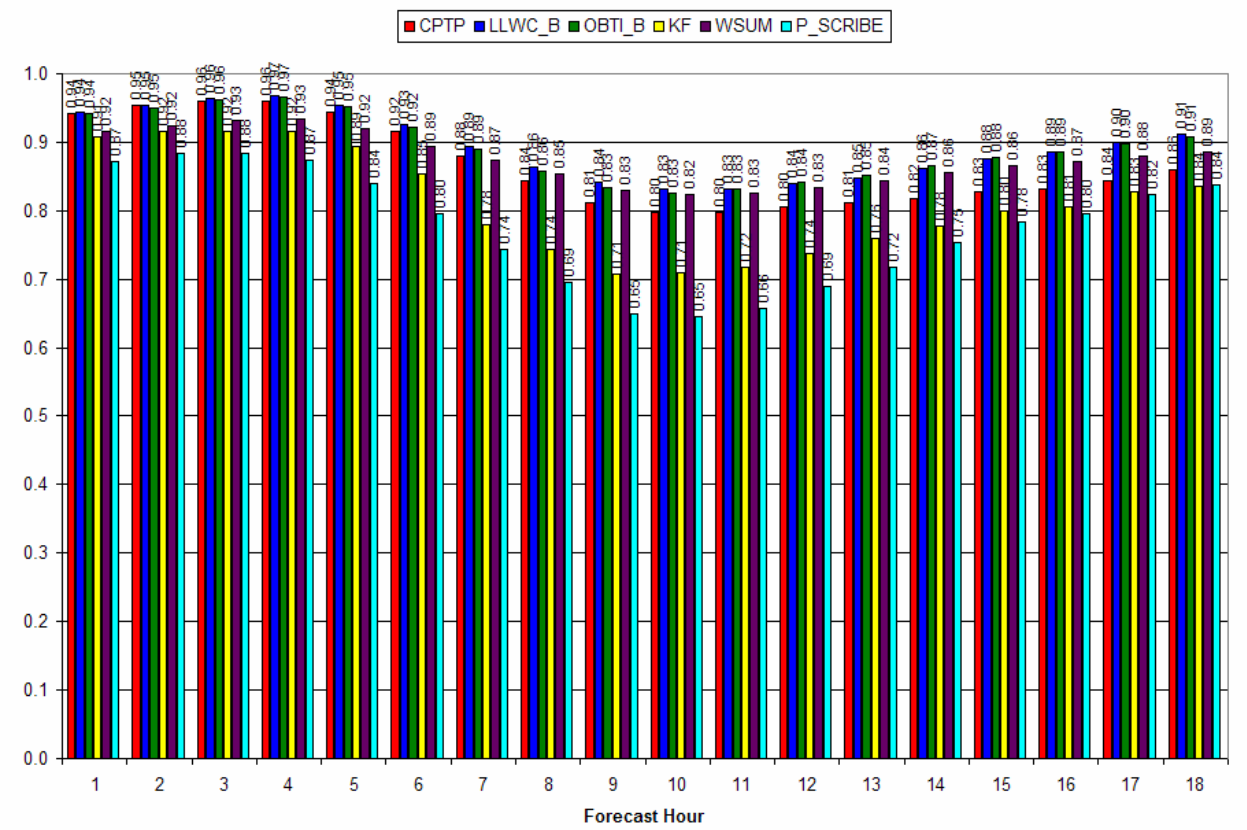

Fig. 19. Objective verification scores for FAR from T+1 h (1300 UTC) to T+18 h (0600 UTC).

5. SUMMARY AND FUTURE WORK

Limited forecast tools or parameters using NWP are currently available to assist Canadian meteorologists in forecasting TI. Work is ongoing at Environment Canada to develop experimental NWP fields targeting TI based largely on model characterization of ABL processes. This is in support of an anticipated move away from point-based forecast production to an area-based system. It is hoped that first-guess TI forecasts can be developed and implemented to aid meteorologists in identifying TI areas.

A number of forecast parameters were developed to consider water vapour availability / depth, low-level convergence / convergence depth, and ABL or near ABL vertical wind shear and buoyancy. Numerous experimental fields were made available to, and evaluated in, PASPC-Edmonton forecast operations via an internal web page during the summer of 2007. These fields were found to be useful for visualizing model behaviour and aiding in identification of potential TI areas. Using observed lightning data from 1 May to 30 September 2007, lightning flashes were associated with parameter values. Apparent relationships between observed lightning and forecast parameter values were used to subjectively define thresholds of selected parameters for forecasting TI.

In 2009, experimental objective forecasts of TI were introduced to PASPC-Edmonton operations and subjectively verified in real-time on the RSD. These forecasts used an ingredients-based approach in their formulation to consider ABL moisture, tropospheric stability, and lift. Existing objective forecasts of convection from NWP data were also verified for comparison. It was noted that for a number of the experimental TI forecasts, areal extent of TI areas was reasonable but verification scores suffered from incorrect model placement or timing of convergence lines and other features. The top forecasts from the subjective verification were LLWC_B, OBTI_B, and a modified version of CPTP with a minimum threshold of 25.

Forecasts of TI from 2009 were objectively verified against observed lightning during the winter of 2009-10. Forecast values at a model grid point were extended to be valid at 24 surrounding grid points to create a 5x5 grid-point tile with dimensions ~ 60x60 km. Standard verification

metrics were used to asses the utility of the experimental TI and existing convection forecasts. Overall verification scores were not impressive, perhaps based in part on using individual lightning flash locations for verifying observations over a very large domain. The detection scores of the experimental TI forecasts improved significantly during the daytime TI period (~1600 UTC to 0200 UTC) when compared to existing forecasts of convection. Results suggest that different choices for TI forecasts may be appropriate based on the forecast period (e.g., morning / night vs. daytime). Certain verification metrics suggest experimental TI forecasts (e.g., LLWC_B, OBTI_B) have superior utility while others favor existing convection forecasts (e.g., CPTP, KF). Our assertion is that detection of rare and hazardous weather elements (lightning in this case) is the most important attribute of TI forecasts. This condition limits the utility of forecasts with superior scores for some selected metric if detection (POD) of the forecast event is poor.

Both subjective and objective verification methods rank the approximation of the SCRIBE-based forecast used here (P_SCRIBE) as the poorest performing forecast for TI and lightning areas. In development of an area-based system, improved first-guess forecasts can likely be attained through the use of the experimental TI or other existing objective convection forecasts considered here.

Objective verification scores suggest there is room for improvement in the formulation of the experimental TI forecasts. The biggest deficiency appears to be the tendency to over-forecast TI areas resulting in poor FAR scores. Future efforts will include more rigorous evaluation of lightning and forecast parameter relationships to define appropriate thresholds for use in TI forecasts. One option is to use classification and regression tree (CART) statistics to define appropriate parameter thresholds. Preliminary work in this regard is underway at the HAL. The application of TI forecasts to ensemble prediction systems is another area to be considered.

As a companion to the TI forecast work, we are in the processing of developing real-time updated TI fields using surface meteorological observations. Forecasts will be re-generated shortly after the current hour using observed surface parameters to recalculate tropospheric stability and ABL characteristics. It is hoped that these products will show utility for thunderstorm nowcasting as TI becomes imminent.

6. APPENDIX

APPENDIX A: A brief description of the objective verification metrics used in this study is given in the table below. Descriptions are taken from, “Forecast Verification: Issues, Methods, and FAQ” by the WWRP/WGNE Joint Working Group on Forecast Verification Research. Available online at: http://www.cawcr.gov.au/projects/verification/.

Name of Metric	Acronym / Abbreviation	Brief Description
Probability of Detection	POD	Fraction of observed “yes” events that were correctly forecast
False Alarm Ratio	FAR	Fraction of the predicted “yes” events that did not actually occur
Critical Success Index	CSI	Correspondence of forecast “yes” events to observed “yes” events
Ratio of POD / FAR	POD/FAR	Simple ratio of POD and FAR to highlight detection vs. false alarm forecasts.
Accuracy	ACC	Fraction of the forecasts that were correct.
Probability of False Detection	POFD	Fraction of the observed “no” events incorrectly forecast as “yes” events.
Hit Rate – Yes	HR(Y)	Fraction of correct “yes” forecasts to all “yes” forecasts.
Hit Rate - No	HR(N)	Fraction of correct “no” forecasts to all “no” forecasts.
Hanssen and Kuipers Discriminant	HK	Skill in separating “yes” events from “no” events.
Heidke Skill Score	HSS	Accuracy of the forecast relative to random chance.
Equitable Threat Score	ETS	Correspondence of forecast “yes” events to observed “yes” events accounting for hits due to chance.
Bias	BIAS	Forecast frequency of “yes” events to observed frequency of “yes” events.
Odds Ratio	ODDS	Odds of “yes” forecast being correct to odds of a “yes” forecast being incorrect.

7. REFERENCES

Buban, M. S., C. L. Ziegler, E. N. Rasmussen, and Y. P. Richardson, 2007: The dryline on 22 May 2002 during IHOP: Ground-radar and in situ data analyses of the dryline and boundary layer evolution. *Mon. Wea. Rev*., **135**, 2473–2505.

Bright, D. R., M. S. Wandishin, R. E. Jewell, and S. J. Weiss, 2005: A physically based parameter for lightning prediction and its calibration in ensemble forecasts. Preprints, *1st Conf. on Meteorological Applications of Lightning Data,* San Diego, CA, Amer. Meteor. Soc.[Available online at: http://ams.confex.com/ams/pdfpapers/84173.pdf]

Côté, J., S Gravel, A. Méthot, A. Patoine, M. Roch, A. Staniforth, 1998: The operational CMC-MRB Global Environmental Multiscale (GEM) model: Part I - Design considerations and formulations. *Mon. Wea. Rev*., **126**, 1373-1395.

Craven, J. P., R. E. Jewell, and H. E. Brooks, 2002: Comparison between observed convective cloud-base heights and lifting condensation level for two different lifted parcels. *Wea. Forecasting*, **17**, 885-890.

Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767-1785.

Crook, N. A., and J. B. Klemp, 2000: Lifting by convergence lines. *J. Atmos. Sci*., **57**, 873-890.

Houston, A. L. and D. Niyogi, 2007: The sensitivity of convective initiation to the lapse rate of the active cloud-bearing layer. *Mon. Wea. Rev*., **135**, 3013-3032.

Kain, J. S., and J.M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784-2802.

Markowski, P., C. Hannon, and E. Rasmussen, 2006: Observations of convection initiation "failure" from the 12 June 2002 IHOP deployment. *Mon. Wea. Rev*., **134**, 375-405.

McCaul, E. W., and C. Cohen, 2002: The impact on simulated storm structure and intensity of variations in the mixed layer and moist layer depths. *Mon. Wea. Rev*., **130**, 1722-1748.

Mueller, C. K., J. W. Wilson, and N. A. Crook, 1993: The utility of sounding and mesonet data to nowcast thunderstorm initiation. *Wea. Forecasting*, **8**, 132-146.

Rotunno, R., J. Klemp, and M. Weisman, 1988: A theory for squall lines. *J. Atmos. Sci*., **45**, 463-485.

Sills, D. M. L., Wilson, J. W., Joe, P. I., Burgess, D. W., Webb, R. M., and N. I. Fox, 2004: The 3 November tornadic event during Sydney 2000: Storm evolution and the role of low-level boundaries. *Wea. Forecasting*, **19**, 22-42.

Sills, D. M. L., and N. M. Taylor, 2008: The Research Support Desk (RSD) initiative at Environment Canada: Linking severe weather researchers and forecasters in a real-time operational setting. *Preprints, 24th AMS Conference on Severe Local Storms,* Savannah, GA, Amer. Meteor. Soc., Paper 9A.1.
[Available online at:
http://ams.confex.com/ams/pdfpapers/142033.pdf]

Sills, D. M. L., 2009: On the MSC forecasters forums and the future role of the human forecaster. *Bull. Amer. Meteor. Soc.*, **90**, 619–627.

Weckwerth, T. M., and R. M. Wakimoto, 1992: The initiation and organization of convective cells atop a cold-air outflow boundary. *Mon. Wea. Rev*., **120**, 2169-2187.

Weckwerth, T. M., J. W. Wilson, and R. M. Wakimoto, 1996: Thermodynamic variability within the convective boundary layer due to horizontal convective rolls. *Mon. Wea. Rev*., **124**, 769-784.

Weckwerth, T. M., D. B. Parsons, S. E. Koch, J. A. Moore, M. A. LeMone, B. B. Demoz, C. Flamant, B. Geerts, J. Wang, and W. F. Feltz, 2004: An overview of the International H_20 Project (IHOP_2002) and some preliminary highlights. *Bull. Amer. Meteor. Soc*., **85**, 253-277.

Weckwerth, T. M., and D. B. Parsons, 2006: A review of convection initiation and motivation for IHOP_2002. *Mon. Wea. Rev*., **134**, 5-22.

Weisman, M. L., Davis, C., Wang, W., Manning, K. W., and J. B. Klemp, 2008: Experiences with 0-36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23**, 407-437.

Wilson, J. W., and W. E. Schreiber, 1986: Initiation of convective storms at radar-observed ABL convergence lines. *Mon. Wea. Rev*., **114**, 2516-2536.

Wilson, J. W., G. B. Foote, N. A. Crook, J. C. Fankhauser, C. G. Wade, J. D. Tuttle, and C. K. Mueller, 1992: The role of boundary-layer convergence zones and horizontal rolls in the initiation of thunderstorms: A case study. *Mon. Wea. Rev*., **120**, 1785-1815.

Wilson, J. W., and R. D. Roberts, 2006: Summary of convective storm initiation and evolution during IHOP: Observational and modeling perspective. *Mon. Wea. Rev.*, **134**, 23-47.

Wilson, L. J., and M. Vallée, 2002: The Canadian updateable model output statistics (UMOS) system: design and development tests. *Wea. Forecasting*, **17**, 206-222.

Ziegler, C. L. and E. N. Rasmussen, 1998: The initiation of moist convection at the dryline: Forecasting issues from a case study perspective. *Wea. Forecasting*, **13**, 1106–1131.