PAPER

Detection of disruptions in the high-β spherical torus NSTX

Nucl. Fusion **53** (2013) 063021 (19pp) doi:10.1088/0029-5515/53/6/063021

Detection of disruptions in the high-β spherical torus NSTX

S.P. Gerhardt[1], D.S. Darrow[1], R.E. Bell[1], B.P. LeBlanc[1], J.E. Menard[1], D. Mueller[1], A.L. Roquemore[1], S.A. Sabbagh[2] and H. Yuh[3]

[1] Princeton Plasma Physics Laboratory, Princeton NJ, 08543, USA
[2] Department of Applied Physics, Columbia University, New York NY, 100027, USA
[3] Nova Photonics, Princeton NJ, 08540, USA

Received 13 December 2012, accepted for publication 22 April 2013
Published 17 May 2013
Online at stacks.iop.org/NF/53/063021

Abstract
This paper describes the prediction of disruptions based on diagnostic data in the high-β spherical torus NSTX (Ono *et al* 2000 *Nucl. Fusion* **40** 557). The disruptive threshold values on many signals are examined. In some cases, raw diagnostic data can be used as a signal for disruption prediction. In others, the deviations of the plasma data from simple models provides the information used to determine the proximity to disruption. However, no single signal or calculation and associated threshold value can form the basis for disruption prediction in NSTX; thresholds that produce an acceptable false-positive rate have too large a missed or late-warning rate, while combinations that produce an acceptable rate of missed or late warnings have an unacceptable false-positive rate. To solve this problem, a novel means of combining multiple threshold tests has been developed. After being properly tuned, this algorithm can produce a false-positive rate of 2.8%, with a late + missed warning rate of 3.7% and thus a total failure rate of 6.5%, when applied to a database of ~2000 disruptions during the I_P flat top collected from three run campaigns. Furthermore, many of these false positives are triggered by near-disruptive magnetohydrodynamic (MHD) events that might indeed be disruptive in larger plasmas with more stored energy. However, the algorithm is less efficient at detecting the MHD event that prompts the disruption process.

(Some figures may appear in colour only in the online journal)

1. Introduction

Tokamak [1] and spherical torus (ST) [2] plasmas can suffer from off-normal termination events called disruptions [3, 4], where rapid losses of energy and plasma current occur in rapid succession; these two events are known as the thermal and current quenches respectively. In most present tokamaks, the primary consequence of a disruption is loss of discharge time. While this will also be true of large reactor-scale facilities, disruptions also have the potential to create significant damage to the facility. Heat loading during the thermal quench [5–8] may melt or ablate significant portions of the first wall or divertor [3, 4, 9, 10]. Mechanical loading by induced eddy currents during the current quench [11–16] can create significant forces on in-vessel structures such as blanket modules [17]. Furthermore, during the current quench, the plasma current can be converted to a beam of runaway electrons [3, 4, 18–26], with potential for severe damage if position control of the beam is lost and it strikes the first wall [27]. Finally, the plasma column often moves upwards or downwards during or proceeding the disruption, and makes contact with the first wall or divertor. When this happens, 'halo' currents [3, 4, 28–36] flowing between the plasma and the wall or divertor structures can lead to large forces on those items. For all these reasons, disruptions must be avoided in these large next-step devices.

The avoidance of disruptions requires a multi-layer approach. Operation scenarios with reduced intrinsic tendency to disrupt should be found, and control techniques must be developed to both maintain those optimal equilibrium characteristics and to prevent the growth of disruptive magnetohydrodynamic (MHD) instabilities. Examples of equilibrium control techniques include control of the strongly shaped plasma boundary [37, 38], the global β_N [39–41], the internal profiles [42–45], or error fields [46, 47]. With regard to instability control, the vertical position instability [48–52] is controlled [50–52] as a matter of routine in all shaped tokamaks. Active control of the resistive wall mode (RWM) [53, 54] has been demonstrated using magnetic mode detection and applied 3D fields [55–60]. Control of the $m/n = 2/1$ neoclassical tearing mode [61, 62] has been demonstrated using gyrotrons to drive currents inside the magnetic island [62–66]. Furthermore, even when instabilities grow large and result in significant modifications of the plasma state, active control 'recovery techniques' can be envisioned, for instance, using electron cyclotron current drive (ECCD) on locked

modes in DIII-D [67] and to avoid high-β disruptions in AUG [68], or electron cyclotron heating (ECH) as demonstrated on AUG [69, 70] and FTU [69] to recover from locked mode and density limit disruptions.

However, on occasion, a disruption is inevitable, and it is necessary to terminate a discharge in a way that minimizes deleterious effects. Methods to accomplish these goals typically involve injecting a large quantity of gas or other material into the discharge, in order to uniformly radiate the plasma thermal energy, promote a current quench before the plasma can drift in the chamber, and if possible suppress runaway electron formation. Technologies for injecting this material include massive gas injection (MGI) [71–79], shell pellets [78, 80], shattered pellets [78, 80, 81], or high-pressure rupture discs [82, 83]. Techniques of this variety have proven useful in reducing the localized thermal loading and halo current forces in tokamaks where they have been used.

A key requirement for these systems is that they be triggered with sufficient warning time to execute their desired tasks. These trigger times can be as late as well into the runaway electron phase, if the mass injection is intended to dissipate the runaway electron beam. However, one of the goals will typically be to eliminate the localized first wall and divertor heating during the vertical displacement event (VDE) or thermal quench. Triggers preceding the beginning of the disruption process will be required for this purpose.

In present machines, that realtime trigger, for either MGI or a rapid rampdown of the stored energy and plasma current, is often provided by a set of 'locked-mode' sensors [75, 84]. These are large magnetic sensor loops mounted to the vessel surface and instrumented for sensitivity to the $n = 1$ radial magnetic field perturbation. These loops can be quite sensitive to the large, quasi-stationary $n = 1$ magnetic islands that typically grow during the disruption process, and thus provide a reasonable basis for triggering mitigation systems.

However, for ITER and likely other future large tokamak or ST facilities, it will be necessary to utilize more than a single diagnostic for triggering mitigation systems. This is because it is not likely that any single diagnostic (including locked-mode detectors) would be capable of detecting all disruptions with sufficient warning time to allow complete mitigation of all phases of the disruption. The primary technique considered so far for combining diagnostic signals for ITER is a neural network [4].

Neural networks are potentially complex functions that mimic biological neural networks. These functions have multiple inputs, and can have multiple outputs; between the inputs and outputs are one or more layers of 'neurons', where the calculations are performed. Neural networks are typically trained, in the sense that a predetermined sample of input and output data are used to determine the optimal values of the coefficients of the network. Neural networks have been used to predict various forms of disruptions on ASDEX Upgrade [85–87], DIII-D [88], ADITYA [89, 90], TEXT [91, 92], JET [86, 93, 94], and JT-60 [95, 96].

For tokamak disruption detection applications, a subset of disrupting discharges is selected for training the network. The efficacy of the network is then determined by testing the predicted disruption alarm on a series of discharges not used to train the detector. Typical input data to the network include a measure of the plasma β (the normalized β, $\beta_N = \frac{\beta_T a B_T}{I_P}(\%\mathrm{mT\,MA^{-1}})$ [97, 98] or poloidal β, $\beta_P = \frac{\langle P_{th}\rangle}{B_P^2/2\mu_0}$), edge safety factor, plasma density (or Greenwald fraction [99, 100]), locked-mode amplitude, input power, radiated power, boundary shape parameters, internal inductance, confinement time, and neutron emission. Some early work also used soft x-ray emission [88–90, 92], high(er) frequency magnetic probes [88–91], or D_α monitors [88–90] though these diagnostics have typically not been used in more recent studies [86, 87, 93, 94, 96].

These techniques have been used with success in disruption detection on many machines. For instance, [94] shows missed and false alarm rates of 23% and 1% respectively from a sample of 86 disrupting pulses and 102 non-disrupting pulses on the JET tokamak. Furthermore, initial cross-machine benchmarking has been completed. Reference [86] shows that a neural network trained on JET data can anticipate 67% of ASDEX Upgrade disruptions within 10 ms, while a neural network trained on ASDEX Upgrade can anticipate 69% of disruptions on JET within 40 ms. While this research represents excellent progress towards the formulations of a disruption detector for ITER, more work is required to understand the extrapolation to that larger device.

In this paper, disruption precursors and detection are studied in the high-β National Spherical Torus Experiment (NSTX), with the goal of determining if disruptions in ST plasmas are generally detectable. Section 2 provides necessary background material on the NSTX device and NSTX disruptions. Section 3 examines the ability of individual diagnostic signals, sometimes coupled to simple models, to predict disruptions. Section 4 describes an algorithm for combining the diagnostic data to form a disruption predictor. A summary and discussion is provided in section 5.

2. NSTX, NSTX disruptions and analysis methods used

NSTX [101] is a medium-scale ST located at Princeton Plasma Physics Laboratory. It has a major radius $R_0 = 0.85$ m, with typical plasma aspect ratios of 1.3–1.55. Typical toroidal fields are in the range $0.35 < B_T(\mathrm{T}) < 0.55$, with plasma currents $0.5 < I_P(\mathrm{MA}) < 1.3$. Plasmas are heated with up to 7 MW of neutral beam injection [102], or ~6 MW of 30 MHz high harmonic fast wave (HHFW) heating [103]. The neutral beams inject their power parallel to the plasma current.

NSTX is equipped with a large number of diagnostics that can be applied to detect disruptions. The electron density and temperature are measured at 60 Hz using a 30-point Thomson scattering diagnostic [104]. The temperature, density, and toroidal rotation of fully stripped carbon are measured with a 51-point charge exchange recombination spectroscopy (CHERS) diagnostic, with time resolution of 10 ms [105]. The central safety factor is estimated in this work from the slope of the pitch angle versus major radius [106], based on the NSTX motional Stark effect (MSE) diagnostic [107].

The magnetic equilibrium is reconstructed using the NSTX installations [108, 109] of the EFIT code [110]. In particular, the reconstructions used in this study are the standard post-shot reconstructions, constrained by

measurements of the coil currents, the magnetic field and flux at various points inside and outside the vacuum chamber, the toroidal plasma and vessel currents [111], the diamagnetic flux, and the electron pressure profile. This set of equilibrium constraints is sufficient to ensure a smooth temporal evolution of profile moments such as the pressure peaking factor and internal inductance, and could in principle be applied to realtime equilibrium reconstruction. Note, however, that because these reconstructions are not constrained by the MSE diagnostic, the central and minimum safety factor evolution from these reconstructions has considerable uncertainty.

The analysis methods used in this paper are 'causal' to the greatest extent possible. That is to say, all digital filters applied to the data use only data from times previous to that under consideration; these include both single-pole low-pass filters [41] and causal median filters. Similarly, signal interpolation is based on previous data only. The single exception to this rule is with regard to the EFIT data, which applies non-causal smoothing to the magnetics data that is used to constrain the reconstruction. Thus, the EFIT data used in the tests should likely be 'shifted' forwards in time by a few ms, or equivalently, a few ms added to the disruption warning times derived from that data. However, the EFIT reconstructions are only used to track the comparatively slow evolution of the plasma, and so this few ms error is not significant. Note that the use of causal low-pass filters introduces an additional consideration into the calculations, in that these filters introduce delays comparable to the filter time constant. These delays can increase the difficulty in detecting rapid disruptions, and it is thus important to minimize the time constants of these filters.

In order to understand the statistical analysis presented in sections 3 and 4, it is useful to examine the dynamics of the disruption process for a typical NSTX disruption. Such data are shown in figure 1. Figure 1(a) illustrates the plasma current evolution for this discharge. The plasma current is maintained by loop voltage feedback at the requested value of 800 kA. There is a small deviation in the plasma current visible at $t_1 = 720$ ms, followed by a much larger drop at $t_2 = 745$ ms (this time will be called t_{IpDev} below, as the time of the first large negative I_{P} deviation). The final current quench occurs at $t_{\mathrm{quench}} = 778$ ms. It is the dynamics during this phase between t_1 and t_{quench} that are of interest.

Figure 1(b) shows the evolution of the plasma density and upper divertor D_α emission. It is clear that the increase of D_α occurs just preceding the time t_1, corresponding to an H $\rightarrow$ L back transition. There is a simultaneous drop in the plasma density, indicative of the loss of the edge transport barrier. Figure 1(c) shows that there are two substantial drops in β_{N}, at times t_1 and t_2. Both of these stored energy drops result in increases in the loop voltage, as the non-inductive currents, proportional to the plasma temperature or β, decrease and the inductive currents must be increased to replace them.

Figures 1(d)–(f) show further signals that will be used below for disruption detection. The neutron emission rate S_{N} is shown in blue in figure 1(d), and illustrates a series of sharp collapses. The neutron emission in NSTX is dominated by beam–target reactions [112], and these collapses are too fast to be the result of rapid fast-ion slowing down. Rather, they

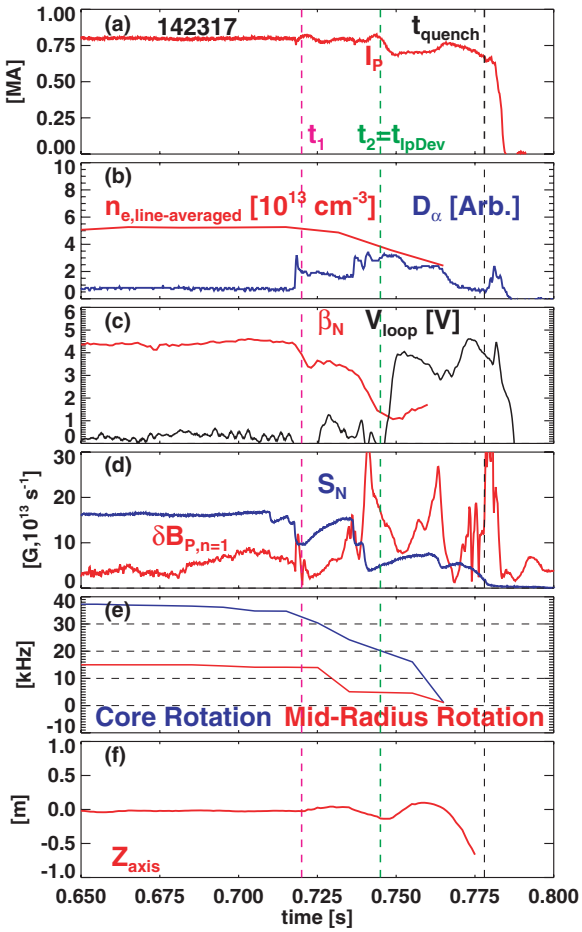

Figure 1. Time evolution of quantities during the disrupting phase of an NSTX discharge. Shown are (a) the plasma current, (b) the line-average density and upper divertor D_α emission, (c) the normalized $\beta(\beta_{\mathrm{N}})$ and the loop voltage (V_{loop}), (d) the neutron emission (S_{N}) and quasi-stationary $n = 1$ poloidal field perturbation ($\delta B_{\mathrm{P},n=1}$), ($e$) the rotation in the plasma core and at the mid-radius and (f) the vertical position of the magnetic axis (Z_{axis}).

are indicative of bursts of fast ion loss [112]. Shown in red in the same figure is the quasi-stationary $n = 1$ poloidal field perturbation, denoted as $\delta B_{\mathrm{P},n=1}$ in this paper, as assessed by an array of internal poloidal field sensors designed to measure slowly rotating low-n perturbations [113]. Signals in these sensors are indicative of the growth of $n = 1$ locked modes [47] and RWMs [47, 109, 113–115]; kink or tearing instabilities rotating at 2–40 kHz are not detected by these detectors. There is strong growth in this signal starting just before t_2, as the plasma develops a strong 3D distortion and locks to the wall. Figure 1(e) shows the rotation in the plasma core and mid-radius. As in [117], the core rotation is defined as the average rotation measured by CHERS channels 4 through 7, spanning $R = 0.99$ to 1.09 m, while the midradius rotation is defined as the average rotation measured by CHERS channels 16–20, spanning $R = 1.27$–1.34 m; these midradius chords are in the vicinity of the $q = 2$ surface for typical NSTX NBI heated discharges. In this case, there is a rapid drop in the mid-radius rotation following time t_1, with the core rotation dropping to zero shortly afterwards. Finally, figure 1(f) shows the plasma vertical position. A series of oscillations in the vertical position grow starting at time t_1, resulting in the plasma eventually impacting the divertor floor.

While the detailed evolution of each disruption can differ, and there may or may not be times equivalent to t_1 and t_2 in each case, much of the phenomenology of this example transfers to other discharges. In particular, there is often a significant duration between the events that initiate the disruption process (at time t_1), and the current quench itself ($t_{\rm quench}$). Sections 3 and 4 will provide a detailed study of how diagnostic signals during this phase can be used to predict disruptions.

This paper will use the phrases false positive, late alarm, and missed alarm, and they are defined as follows. A false positive occurs when the disruption alarm is declared more than 0.3 s in advance of the disruption start time $t_{\rm dis}$; this duration corresponds to ~10 energy confinement times, or ~1 current redistribution time. Note that the declaration of a false positive does not imply that the discharge did not disrupt, only that the disruption warning anticipated the disruption by too great a time for there to be any relationship between the warning and the disruption. A late warning is defined to occur if the alarm is declared after $t_{\rm dis} - 0.01$. Both the 0.3 and 0.01 s definitions (for false positives and late warnings) are somewhat arbitrary, and the figures below show the distribution of warning times to determine how changing these two definitions would affect the results. A missed warning occurs when the disruption alarm is not declared at any point during the disrupting discharge.

For most examples in this paper, the disruption time ($t_{\rm dis}$) will be defined to be equal to $t_{\rm quench}$, determined from on the $I_{\rm P}$ waveform using an automated method as described in [14]. If an $I_{\rm P}$ spike, generally indicative of the thermal quench [1] is present in the $I_{\rm P}$ waveform, then $t_{\rm quench}$ is defined as the start of the spike. If no spike is found, then $t_{\rm quench}$ is defined as 2 ms before the final collapse of the plasma current. This definition of $t_{\rm quench}$ is chosen because it can be readily evaluated for all NSTX discharges; there are not any indicators of the thermal quench timing that can be so readily evaluated. However, this definition, based on the leading edge of the $I_{\rm P}$ spike or just preceding the current decrease, is generally coincident with the thermal quench. An additional relevant time is that when $I_{\rm P}$ first deviates meaningfully from its requested value ($t_{\rm IpDev}$) [117]; this time is typically, but not always, before $t_{\rm quench}$, due to the often-significant pre-disruption energy and current losses. Hence, in section 4.5, the definition $t_{\rm dis} = \min(t_{\rm quench}, t_{\rm IpDev})$ will be used.

The present analysis uses data from the 2008, 2009, and 2010 run campaigns, during which many meaningful changes to the facility occurred. At the beginning of the 2008 campaign, neither feedback-based RWM control and error field correction, nor lithium conditioning of the plasma-facing components (PFCs) [118, 119], were commonly used. However, over the course of that campaign and throughout the 2009 run, those techniques became routine. The 2010 campaign saw a major change of the PFCs, in the form of the liquid lithium divertor (LLD) [120]. This change not only replaced a fraction of the graphite PFCs with a porous molybdenum surface, it resulted in a substantial increase in the fraction of discharges with lower-triangularity, so that the outer strikepoint would be in the vicinity of LLD. This is to be contrasted to previous campaigns, where higher-triangularity discharges were generally preferred due to the improved performance of discharges with strong shaping [121–123].

3. Individual disruption precursors

As indicated in the previous section, there are many potential precursors present before NSTX disruptions. This section examines individual precursor signals for their suitability as disruption predictors, including discussion of some underlying physics. All disruptions considered in this section occurred during the $I_{\rm P}$ flat-top phase of the discharge.

3.1. Magnetic measurements

A key set of measurements for disruption prediction involve magnetic perturbations associated with the global MHD mode distortion. In this sub-section, measured $n = 0$ and $n = 1$ perturbations, as well as deviations in the plasma current, are assessed as disruption indicators.

As noted above, NSTX has routine measurements of the quasi-stationary $n = 1$ poloidal and radial field perturbations, indicative of the growth of $n = 1$ locked modes and RWMs. This information comes from analysis of the signals from 24 poloidal field sensors and 24 radial field sensors [113], all mounted inside the vacuum chamber. The poloidal field sensors are mounted in thin-wall stainless steel boxes attached to the passive plates, while the radial field sensors are mounted directly in front of the thick copper passive plates. As a consequence, the radial field sensors are significantly slower to detect the mode, but are also less susceptible to noise. A simplified version of the sensor analysis has been executed for RWM control in the majority of H-mode discharges starting with the 2008 run campaign; this analysis is archived after each discharge. A more sophisticated level of sensor analysis, which includes correction of various sensor non-ideal effects, is available for nearly all discharges taken in the device during the operations period under consideration. The analysis presented here is based on the archived realtime analysis, in order to verify that such signals and processing are appropriate for disruption detection.

Figures 2(*a*) and (*b*) show data illustrating the use of the poloidal field and radial field sensors, respectively, for disruption detection. In particular, these figures show histograms of the time between the sensor data achieving the condition indicated by the caption, and the current quench occurring ($t_{\rm dis} = t_{\rm quench}$, see section 2). One dashed vertical line is drawn at 0 ms warning time, and occurrences to the left of this line + 10 ms are late warnings by the criterion indicated towards the end of section 2. A second vertical line is drawn at a warning time of 300 ms, and occurrences to the right of this line are considered false positives. Note that as the thresholds for declaring a warning increase, an ever larger number of discharges will not have these threshold conditions met; this is indicated by the decreasing total area of the histograms for large threshold values.

Returning to the magnetic sensor data, it is clear that defining thresholds of $\delta B/B_{\rm T}$ of $\sim 10^{-3}$ result in only a very marginal disruption indicator for the $B_{\rm P}$ sensors, and have absolutely no predictive ability for the $B_{\rm R}$ sensors. Increasing the threshold to 2.5×10^{-3} results in reasonable disruption indicators for both magnetic field components, with minimal false positives, but an increasing fraction of late warnings. Increasing the threshold further to 5×10^{-3} results in complete

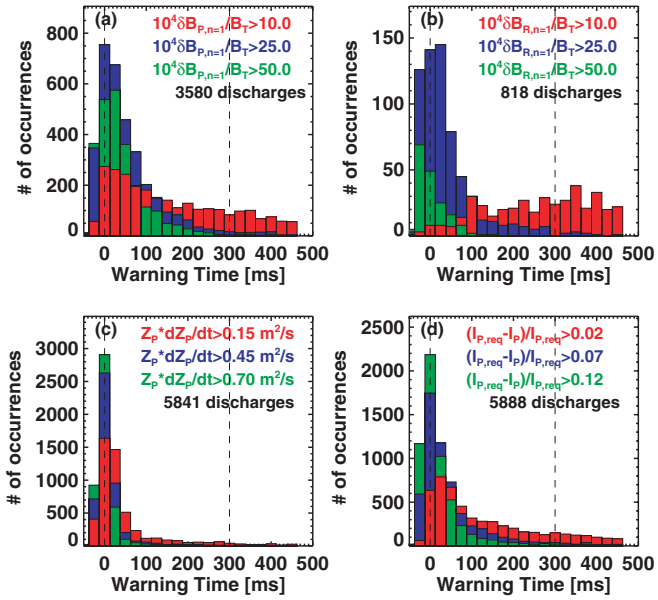

Figure 2. Histograms of disruption warning times, based on (*a*) the normalized quasi-stationary $n = 1$ poloidal field perturbation, (*b*) the normalized quasi-stationary $n = 1$ B_R field perturbation, (*c*) the quantity $Z_P \cdot dZ_P/dt$, and (*d*) the fractional deviation of the plasma current from the requested value.

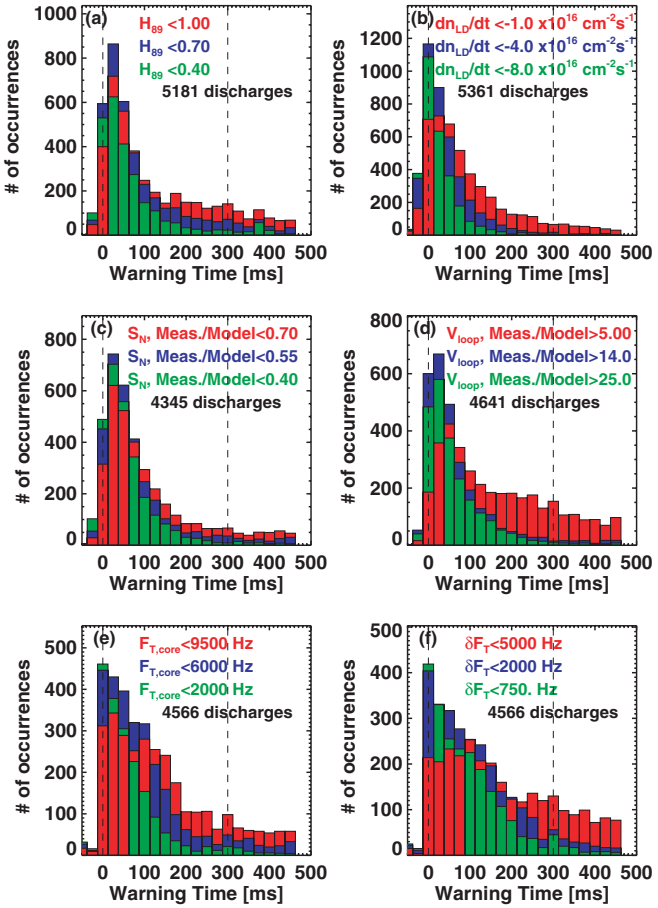

Figure 3. Histograms of disruption warning times, based on (*a*) the global energy confinement level, (*b*) transients in the line-density evolution, (*c*) the neutron emission, (*d*) the loop voltage, (*e*) the core toroidal rotation, and (*f*) the differential rotation between the core and mid-radius.

elimination of false positives, but also a large increase in the fraction of late warnings and a significant reduction in the total number of discharges where the warning is declared.

It is quite common for NSTX plasmas to develop vertical instabilities at some time during the disruption process. These have been observed both during the flat-top phase of high-elongation discharges, and during the phase after a locked mode or RWM drives the plasma out of H-mode [117]. Hence, observing vertical motion can be a useful indicator of proximity to disruption.

For the purposes of this section, the quantity $Z_P \cdot dZ_P/dt$ will be used as an indicator of proximity to disruption. It has the advantage of being large when the plasma is above the midplane and moving upwards, or below the midplane and moving downwards. The two terms are computed using poloidal flux loops above and below the midplane on the outboard side of the plasma. The plasma position is estimated based on the difference in flux measured by the loops, $Z_P I_P = 1.37(\psi_{Upper} - \psi_{Lower})$, while the velocity is estimated from the voltage measured on the flux loops, $\frac{d(Z_P I_P)}{dt} = 1.37(V_{Upper} - V_{Lower})$; I_P is the plasma current measured in MA for these formulae, with Z_P the plasma centroid in the vertical direction, in meters. Note that the flux here is simply the integral of the voltage, as computed by hardware analogue integrators. Figure 2(*c*) shows the results of using this quantity $Z_P \cdot dZ_P/dt$ as an indicator of proximity to disruption. Taking a threshold of 0.15 m^2 s^{-1} results in a reasonably good indicator of disruption probability, with minimal false positives. Larger values eliminate the false positives entirely, though at the expense of an increasing number of late and missed warnings. The need for more sophisticated vertical stability calculations is discussed in section 5.

Finally, as noted in section 2, the pre-disruption activity that reduces the stored energy also reduces the non-inductive current sources, and often results in a drop in the plasma current. The magnitude of this drop, when normalized to the requested current, is an excellent indicator of imminent disruptions. Figure 2(*d*) shows that plasma current deviations of order 2% are often, but not always, associated with proximity to disruption. However, deviations of order 7% and larger are almost always rapidly followed by a disruption.

3.2. *Confinement measurements*

A significant loss of confinement can indicate a degradation of the plasma state, with associated increase in the likelihood of disruption. In this section, it will be shown that loss of fast particles, thermal density or energy, or toroidal angular momentum, or a significant increase in the loop voltage, can all indicate proximity to disruption.

The first confinement indicator selected is the global energy confinement τ_E, normalized by the ITER-89 L-mode scaling expression (τ_{89}) [124]; this quantity is denoted $H_{89} = \tau_E/\tau_{89}$. This confinement expression is evaluated more easily than the ITER-98 H-mode expression [125], as it does not require that the fast particle energy be subtracted from the total energy. Figure 3(*a*) shows that confinement levels dropping to $H_{89} = 1.0$ can often be indicative of proximity to disruption. However, confinement dropping beneath $H_{89} = 0.5$ predicts disruption in almost all cases.

Just as with the thermal confinement, a significant drop in particle confinement is often indicative of disruption imminence. This is shown in figure 3(*b*), where the time derivative of the line-density is used to predict disruptions. Small drops in the line density ($dn_{e,LD}/dt < -10^{16}\,cm^{-2}\,s^{-1}$) can sometimes predict disruptions, though setting that threshold would result in many false positives. Larger negative transients in the line density are typically associated with large-scale rearrangements of the equilibrium, for instance, an H $\rightarrow$ L back transition, and can be an excellent indicator that a disruption is approaching.

As noted before, it has been shown in [112, 117] that MHD events can lead to rapid drops in the neutron rate, indicative of a significant loss of fast ions; figure 1(*d*) shows an example where there are large drops in the neutron emission before the disruption. The utility of this information for disruption prediction is increased if there is an accurate, rapidly evaluated model that can be used to predict the neutron emission. Here, a slowing down model is applied for this purpose, using only the measured electron density and temperature from Thomson scattering and a line-average Z_{eff} from a visible bremsstrahlung measurement. In figure 3(*c*) the ratio of the measured emission to that predicted by the model is used as an indicator of the proximity to disruptions.

Setting a measurement to model ratio of 0.7 results in a histogram of warning times that is peaked towards small warning times, but there remain a large number of false positives. These false positives could be due to errors in the diagnostic data (Z_{eff} in particular), or to neutron emission drops due to bursting Alfvenic modes [126, 127] that are not captured in the simple slowing-down model. However, when the ratio of measured to predicted neutrons drops beneath 0.4, the analysis shows that a disruption almost always rapidly follows.

As noted in sections 2 and 3.1, when there is a rapid loss of plasma energy, density, and fast particle content, it follows that the non-inductive current drive sources will be dramatically reduced as well. It then follows that in order to maintain the plasma current, the loop voltage must increase. Hence, increases in the loop voltage can be an indicator of disruption proximity.

As with the neutron emission, the loop voltage data can be better interpreted if there is a reasonable model of what the loop voltage should be. Such a model was used in [128] to project operating points for NSTX-U, and is described here briefly. The plasma current, toroidal field, density, heating power, and geometry are used to evaluate the ITER-98pb(y,2) scaling expression [125] to predict the stored energy, volume average pressure $\langle P_{th}\rangle$, the poloidal beta β_P, and the average temperature $\langle T_e\rangle$. Note that this expression for the global energy confinement has been shown to predict the energy confinement in NSTX H-mode discharges [123] with reasonable accuracy when lithium conditioning [118, 119] is used, as was the case in the majority of discharges used in this study. These data are then used to estimate the bootstrap current [129–133] and neutral beam current [134–136] from simple scaling expressions. The bootstrap fraction is calculated as $f_{BS} = C_{BS}\sqrt{\varepsilon}\beta_P$, where the coefficient $C_{BS} = 0.4$ is calibrated from detailed calculations of the bootstrap current using the Sauter model [132] in TRANSP [137]. The neutral beam current drive is estimated as $I_{NBCD} = C_{NBCD}\langle T_e\rangle P_{inj}/\bar{n}_e$, and is calibrated against NUBEAM [138] calculations within TRANSP. Once these non-inductive sources are calculated, the inductive currents are calculated as those required to produce the total current ($I_{inductive} = I_P(1 - f_{BS}) - I_{NBCD}$), and the loop voltage estimated based on the expected neoclassical resistivity and inductive current level.

The quantity to be used for disruption prediction is the ratio of the measured to modelled loop voltage. One might think that ratios of order of 3–5 would be indicative of the plasma deviating from the high-performance state. However, there are significant fluctuations in the loop voltage, with the instantaneous value often exceeding the mean by factors of 5 or more. While these transients could be smoothed using one of the available causal filters, the time-scale required for those filters would introduce an unacceptable delay in the response of the test. Hence, it has been found desirable to use very large thresholds on the ratio of the measured to modelled loop voltage when constructing these disruption tests. In particular, as shown in figure 3(*d*), declaring a disruption warning when the measurement exceeds the model by a factor of 5 would result in many false-positives. It is only when the threshold for declaring a warning is placed in the vicinity of $\sim$20 that the false-positive count can be reduced sufficiently.

Finally, the damping of the plasma rotation can form a basis for detecting impending disruptions. This comes about from two observations. First, when the rotation is reduced, it can lead to a reduction in the RWM stability of the system [47, 115] (more correctly, the dependence of RWM stability on rotation can be non-monotonic [139–143], but experimentally, non-resonant magnetic braking is typically used to generate these modes in NSTX). Hence, it is important to apply error field correction techniques that maintain the rotation [47, 144], while monitoring the rotation for drops that may render the plasma more susceptible to these modes. Secondly, the core rotating $n = 1$ modes that often occur in NSTX H-mode discharges [117, 123, 127, 145–147] have a tendency to first reduce the core rotation compared to the edge value, locking the plasma into a rigidly rotating system, followed by a rapid damping of the total plasma rotation. Hence, an observed reduction in the rotation can be indicative of these modes growing large enough to impact the plasma performance, potentially leading to disruption.

Two rotation metrics for disruption detection are considered here: the value of the core rotation speed in figure 3(*e*), and the differential rotation between the core and edge ($\delta F_T = F_{T,core} - F_{T,mid\text{-}radius}$) in figure 3(*f*). When the core rotation drops beneath $\sim$10 kHz, the probability of disruption increases, though using this value as a threshold would result in many false positives. On the other hand, core rotation values beneath 2 kHz almost always lead to a disruption. Similarly, while a differential rotation threshold of 2 kHz would result in many false positives, a threshold of $\sim$750 Hz would predict many disruptions without a large false-positive rate.

3.3. Equilibrium measurements

The final set of measurements we have examined for purposes of disruption detection are those derived from equilibrium

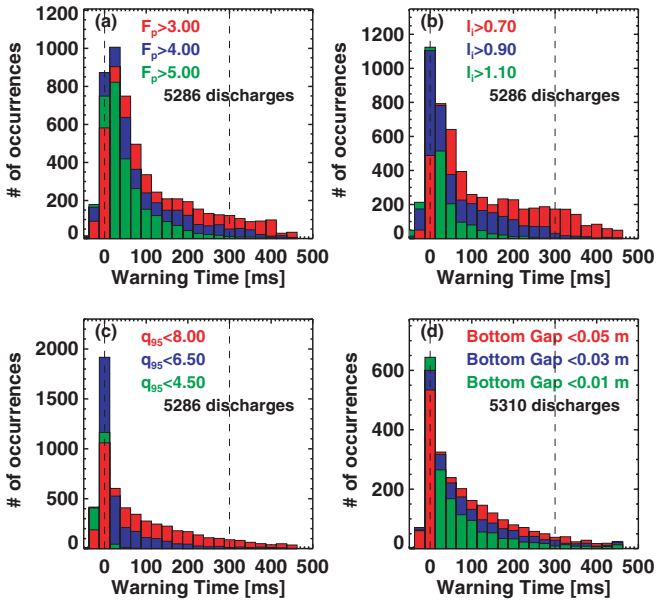

Figure 4. Histograms of disruption warning times, based on (*a*) the pressure peaking factor $F_{\mathrm{P}} = p_0/\langle p\rangle$, (*b*) the internal inductance l_{i}, (*c*) q_{95}, and (*d*) the lower plasma–wall gap.

reconstruction. Example data from this category are shown in figure 4, where threshold tests on the MHD pressure peaking factor ($F_{\mathrm{P}} = p_0/\langle p\rangle$), internal inductance ($l_{\mathrm{i}} = I_{\mathrm{inductive}} = I_{\mathrm{P}}(1 - f_{\mathrm{BS}}) - I_{\mathrm{NBCD}}$), safety factor at the 95% poloidal flux surface (q_{95}), and bottom plasma–wall gap are used for prediction.

With regard to the profile-parameters in the top row, it is clear that increased peaking of either the pressure or current profile can be used as an indicator of disruption proximity. Figure 4(*a*) shows that a pressure peaking threshold of 3.0 results in a large number of false-positives. However, increasing the threshold to 4.0 and then 5.0 reduces the false-positive rate, ultimately resulting in a function peaking with warning times of 20–40 ms. A similar result occurs for the internal inductance l_{i} in figure 4(*b*). Here, a threshold of 0.7 is clearly not indicative of disruption imminence. However, increasing the limit to 0.9 and 1.1 results in ever improving predicative ability for disruptions. Note that both of these trends are consistent with the disruptivity analysis in [117]; see that paper for additional discussion of the underlying physics.

The final equilibrium parameters under study as a disruption indicator are the edge safety factor (in figure 4(*c*)) and bottom plasma wall gap (in figure 4(*d*)). Both of these quantities are sensitive indicators of the loss of plasma vertical position control. For instance, small or zero values of the bottom gap may be indicative of the plasma being in the late phase of a downward-going VDE. Similarly, a drop in the edge-q accompanies the late phase of a VDE as the plasma cross-section shrinks at approximately constant plasma current.

However, both parameters may capture other characteristics that are indicative of disruption tendency. With regard to the bottom gap, small values at fixed plasma vertical position can be related to loss of X-point position control. Having the lower X-point come too close to the divertor floor can result in loss of H-mode, which is generally disruptive in NSTX due to the lower pressure limits with L-mode profiles. With regard to the value of q_{95}, it is clear that the tendency to disrupt increases in NSTX at low q_{95}, even when only centred plasmas are considered. This was shown in [117] to not be due to the low-q limit at $q^* \sim 1.7$ [148], but rather a series of operational issues that tend to increase the likelihood of disruption.

With these caveats in mind, it is clear that setting a q_{95} limit greater than 8, or a bottom gap limit greater than 4 cm, results in an unacceptable level of false positives. However, reducing the thresholds to a bottom gap of 1–2 cm, or q_{95} values in the range of 5–6, eliminates most false positives and can provide an indication that the plasma has entered a regime of increasing probability of disruption.

4. Combinations of disruption precursors

From the above discussion, it is clear that no single diagnostic can produce a reasonable disruption proximity indicator: setting low values of the various thresholds can produce a predictor with a minimal late-warning rate. However, this will typically result in an unacceptable number of false positives. Eliminating the false positives by raising the thresholds results in an increase in the number of late or missed warnings. In order to eliminate this problem, it is necessary to combine the threshold tests in some fashion. This section describes a new algorithm for implementing that combination. In particular, section 4.1 describes the scheme in general terms, section 4.2 describes how the coefficients of the detection algorithm are determined, section 4.3 presents results showing the efficacy of the algorithm, section 4.4 examines the causes of late warnings and false positives, section 4.5 examines the stability of the algorithm across multiple run years, and section 4.6 examines the detectability of the disruption-initiating event. Finally, section 4.7 describes the results of applying this detector to disruptions that occur during the I_{P} rampdown.

4.1. New scheme developed for NSTX data

This paper proposes a scheme somewhat intermediate between disruption detection based on a single diagnostic and threshold on the one hand, and neural networks on the other. The algorithm in this scheme is described below.

To define the algorithm, the following two steps are executed.

- A series of ~17 threshold tests such as those described above is defined.
- For each of these tests, a unique 'point' value is assigned to each of a set of threshold levels. For instance, for a three-level scheme, corresponding to three threshold levels assigned for each test, a typical point assignment is as given in table 1, while the appendix contains the complete set of tests and point assignments. The means of determining these point values will be described in section 4.2.

Then, at each time step during a discharge, the following steps are executed.

(1) Each of the threshold tests is executed, and the number of points for each test is evaluated. For instance, from the first line of table 1, a single point would be awarded if $\delta B_{\mathrm{P},n=1}/B_{\mathrm{T}}$ exceeded 0.0023, two points if it exceeded 0.004, and three points if it exceeded 0.0072.

Table 1. Subset of the point assignments for a three-level test. The total set of point assignments for a three-level test can be found in the appendix. FPR stands for false-positive rates, and the connections between false positive rates and these point assignments will be discussed in section 4.2.

Test	1 pt (5% FPR)	2 pts (2% FPR)	3 pts (0.5% FPR)
$\delta B_{P,n=1}/B_T >$	0.0023	0.004	0.0072
$S_N <$, measurement/model	0.49	0.39	0.27
$(I_{P,req} - I_P)/(I_{P,req}) >$	0.08	0.12	0.19
$F_{T,mid\text{-}radius} - F_{T,core} <$ (kHz)	1.4	*	*

(2) The points from the individual tests are totalled, to form the 'aggregate' point total.

(3) A disruption warning is declared if the aggregate point total exceeds a pre-defined threshold value.

Additionally, it is sometimes observed that a diagnostic signal will transiently cross a threshold value before moving back to the nominally non-disruptive level. If this occurs while other diagnostics are achieving levels indicative of disruption, it would be beneficial to have some memory that the test indicated, at least transiently, an increased probability to disruption. Hence, whenever a diagnostic moves from having violated the lowest threshold level of a test to no longer violating that threshold, the point total associated with the lowest threshold level is retained for a fixed duration called the 'reset window'.

As noted in step 3 above, the disruption warning is declared when the aggregate point total exceeds some fixed pre-defined value. There are multiple ways that this value of the aggregate point total can be achieved. On the one hand, if many individual tests are producing low point totals, a large aggregate total can be accrued. This corresponds to the case where many tests indicate that the plasma is entering a disruptive state, though no single one is conclusive. On the other hand, if a single test produces a large point total, then the one diagnostic can indicate that a disruption may be imminent.

Finally, in addition to the single measurement tests described above, the code has the ability to handle 'dual tests' of the form, 'if quantity a exceeds a given value while quantity b is less than a specified value, then point total c shall be added to the aggregate total.'

At present, the only test of this type is based on the detection of $f < 40$ kHz rotating odd-n poloidal field perturbations at the vessel wall just beneath the outboard midplane. This signal is determined by subtracting the numerically integrated signals between two rapidly sampled magnetic field sensors mounted with 180° toroidal separation from the vessel wall. This signal is typically dominated by $n = 1$ perturbations, and is thus sensitive to $n = 1$ kink/tearing modes that subsequently lock to the wall and cause disruption. For the present example, the thresholds are formulated so that if the frequency drops below 2 kHz while the mode amplitude exceeds 7 G, the point total is incremented. This actual point total is based on that associated with a 1.5% false-positive rate; see discussion below for motivation on the connections between point totals and false-positive rates.

4.2. Determination of point assignments

It is clear from the discussion of section 4.1 that determining the threshold values associated with each point assignment and test is the key step in setting up this algorithm. The method used to date has been to assign thresholds based on values that result in a given false-positive rate.

The rationale behind this choice is illustrated in figure 5. Each frame shows the rates of late warnings (blue), false positives (red), good warnings (green), and total triggers (black), as a function of the threshold level for a single diagnostic signal or calculation. Here, the rate is defined as the number of discharges that have a certain feature (false positive, etc), normalized to the total number of discharges in the dataset. The signals considered, and associated histograms above, are (a) the normalized quasi-stationary $n = 1$ B_P perturbation as in figure 2(a), (b) the pressure peaking as in figure 4(a), (c) the line-averaged density transient as in figure 3(b), and (d) the confinement multiplier H_{89} as in figure 3(a).

The four frames show similar trends in the threshold levels, though there is a reflection of the x-axis based on whether the test is based on the diagnostic signal being bigger or smaller that a given threshold. Consider first declaring a warning when the signal is larger than a given value, as in figures 5(a) and (b). Small threshold values have a high rate of false positives but a low rate of late warnings. Increasing the threshold results in a decrease of the false-positive rate, and an initial increase in the rate of good detection. However, for larger vales of the threshold, it is common for the good detection rate to decrease. This is because many disruptions will not produce diagnostic signals that large, and hence no trigger, either early or late, occurs. This can be seen most clearly in the black curve, which shows the sum of the rates of good detection, late warning, and false positives, or equivalently, the fraction of discharges with any trigger at all; this curve always decreases as the threshold level increases for these tests. As a consequence of these observations, of the three curves corresponding to late warning, good detection, and false-positive, only the false-positive curve is a monotonic function of the threshold value. It is for this reason that the false positive rate is used to determine the threshold values in the tests described in section 4.1. Note also that all of the statements above apply to the tests in frames (c) and (d), where the warning is declared if the value is less than a given value, as long as 'increasing threshold level' is replaced with 'decreasing threshold level'.

Thus, point levels and thresholds for the individual tests are determined based on assigning a number of points to a given false-positive rate. For the three-level test described in table 1, the first row indicates that 1 point is associated with the threshold providing a false-positive rate of 2%, 2 points are associated with a false-positive rate of 1%, and 3 points are associated with a false-positive rate of 0.5%. For the six-level test, the pairs (points, false positive %) are [1,10], [2,5], [3,2], [4,1], [5,0.5] and [7,0.2]. This method provides a deterministic mechanism for the determination of the point totals, and, as shown in the next section, provides a reasonable compromise between false positives and late warnings.

Finally, we note that some diagnostic tests never achieve the lowest values of false-positive rate indicated in the previous paragraph. An example of this is the test on the differential rotation $F_{T,mid\text{-}radius} - F_{T,core}$ in the bottom row of table 1,

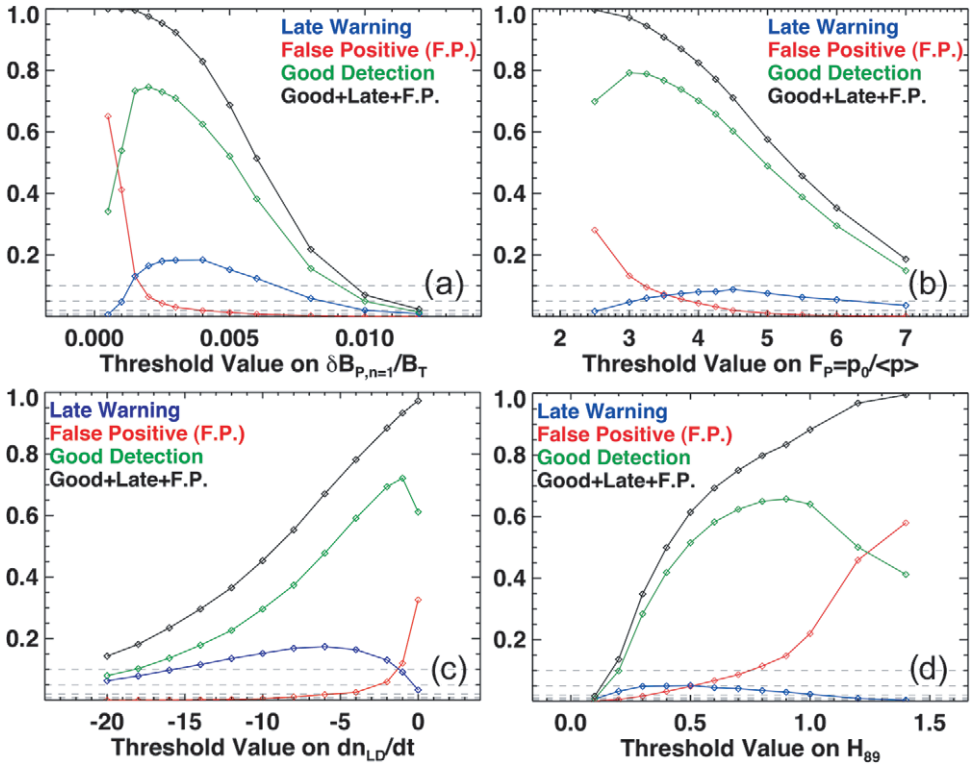

Figure 5. Late warnings, false positives, good detections, and total triggering rate, as a function of the threshold level, for various diagnostic tests. Quantities include (*a*) the normalized quasi-stationary B_P, $n = 1$ signal, (*b*) the pressure peaking, (*c*) the line-average density transient, in units of 10^{16} cm^{-2} s^{-1}, and (*d*) the confinement level.

which has a false-positive rate between 2% and 5% even at the lowest thresholds. In cases like this, only the lower point values associated with the higher false-positive rates are allowed. This provides an automatic de-weighting of those tests with intrinsically high false-positive rates.

4.3. Detector results

With the basis for determining individual point assignments above, this disruption detector is a surprisingly sensitive indicator of proximity to disruption. This will first be shown by analysing the time evolution of the aggregate point total for three individual disrupting discharges in figure 6. The analysis will then be done on a statistical basis for a three-year database of 2026 disrupting plasmas.

The left-most column of figure 6 shows the waveforms for a case where an early MHD mode slows the toroidal rotation and locks to the wall, leading to a disruption; this case was examined in section 6 of [117]. This flavour of disruption is comparatively easy to detect. In particular, as shown in figure 6(*a*) the large-amplitude rotating mode that provides the rotation braking is easily detected with the wall-mounted magnetic field diagnostic sensors when processed to detect the rotating odd-*n* perturbation, as described at the end of section 4.1. In this case, the mode locking is indicated by the rapid drop in odd-*n* amplitude at 0.32 s. The rotation braking itself is also easily detected, as shown in the rapid rotation damping after $t = 0.23$ s in frame (*c*). Once the mode locks, many further diagnostic indicators are clear. A large stationary $n = 1$ B_P perturbation, denoted $\delta B_{P,n=1}$ and normalized to the toroidal field strength in the figure, grows larger, and oscillations in the vertical position begin to grow. The normalized β drops rapidly, and the pressure peaking increases. The loop voltage becomes quite large; as noted above, the cold, L-mode plasma, with reduced non-inductive current sources, requires a significantly larger voltage to maintain the plasma current. However, despite this increased loop voltage, there is still a considerable drop in the plasma current. These trends, when combined with the threshold tests described in section 4.1, lead to a rapid increase in the aggregate point total shown in frame (*d*), in this case exceeding a value of 30 well before the final current quench.

The central column of figure 6 shows an example where a hot-plasma VDE initiates the plasma disruption. This discharge was designed [123] to prototype discharges for NSTX-Upgrade [128], where there is an increase in the passive VDE growth rate [149] due to the larger aspect ratio ($A = 1.67$). In this case, a small increase in l_i resulted in the growth of the VDE. There are essentially no predictors of the disruption in the transport-like variables F_P, β_N, or the loop voltage in frame (*f*), and the rotation in frame (*g*) is unchanged. However, frame (*e*) shows that there is a strong oscillation in the quantity $Z_P \cdot dZ_p/dt$ (and the value of dZ_P/dt as well, not shown in the figure). The warning levels based on these quantities causes the aggregate point total to increase above ten before the current quench.

Finally, the third column shows the results for an RWM disruption. The evidence of the RWM is provided in frame (*i*), where the $n = 1$ perturbation begins to grow rapidly at $t = 0.748$ s. The same frame shows that the plasma current begins to rapidly droop at this time, though at a rate too slow to be considered a disruption; the actual current quench only occurs at $t = 0.787$ s, when the current has dropped by about half. There is also strong vertical motion following

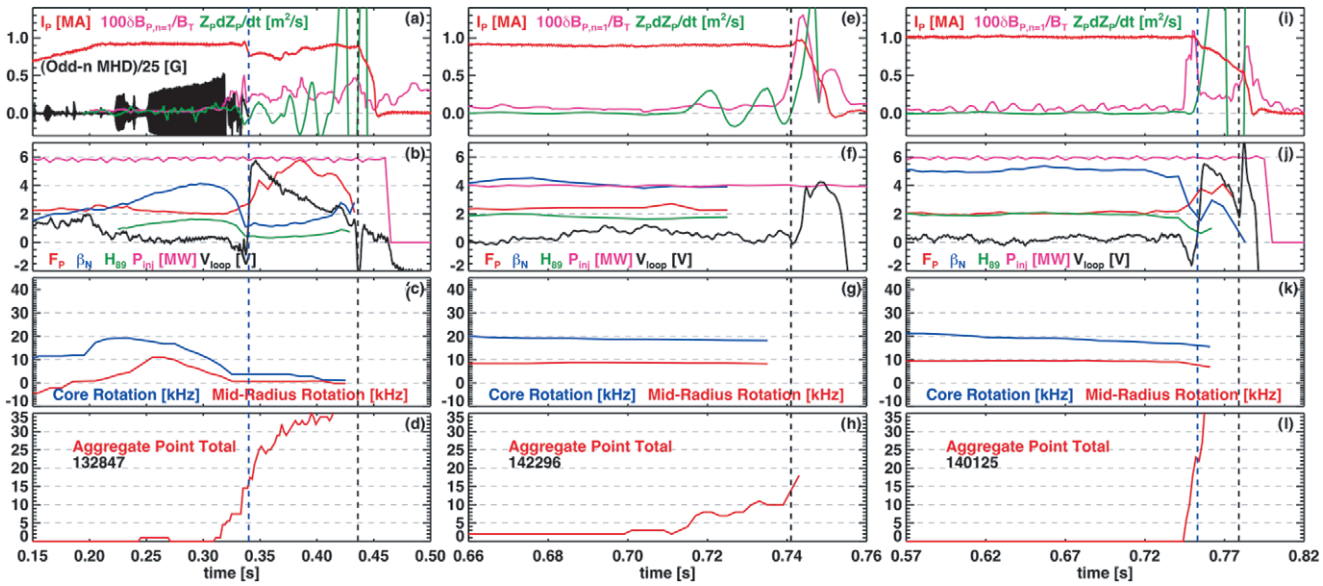

Figure 6. Example time evolution of disrupting plasma. The first row of each column shows the plasma current, normalized $n = 1$ quasi-stationary B_P perturbation ($\delta B_{P,n=1}/B_T$), the rotating odd-n B_P perturbation, and vertical motion indicator $Z_P \cdot dZ_P/dt$. The second row of each column illustrates the pressure peaking (F_P), normalized beta (β_N), confinement enhancement relative to ITER-89 scaling (H_{89}), the injected power (P_{inj}), and the loop voltage (V_{loop}). The third row shows the core and mid-radius toroidal rotation frequency, while the bottom row shows the aggregate point total based in the six-level disruption warning rule. The time of first I_P deviation (t_{IpDev}) is shown as a blue vertical line in the left and right columns, while the black line indicates the time of the current quench initiation (t_{quench} from section 2).

the mode onset, as indicated by the trace of $Z_P \cdot dZ_p/dt$. Many other diagnostic signals in frame (j) register the effect of the mode, including a rapid evolution in the loop voltage, increase in the pressure peaking, and drop in confinement. This particular case is too rapid for there to be any observation of strong rotation damping preceding the mode, and the CHERS diagnostic, providing the data illustrated in frame (k), is unable to resolve the likely rapid drop in rotation following the large mode onset. Nevertheless, the observed signals are sufficient to trigger an extremely rapid rise in the aggregate point total in frame (l), well before the final current quench.

These results are summarized on a statistical basis in figure 7, which shows in frames (a) and (b) the warning time statistics for both three- and six-level tests; each frame shows statistics based on two thresholds of the aggregate point total above which a disruption warning is declared. This is from a database of 2026 disruptions during the I_P flat-top of beam-heated discharges. Furthermore, only discharges with maximum stored energies greater than 20 kJ, plasma currents greater than 680 kA, and with disruption current quench rates less than -2.2×10^7 A s^{-1} are allowed in the dataset.

The three-level test in figure 7(a) shows a 21.5% 'failure rate' for warning threshold of 2 points; as per table A1, this warning threshold corresponds to any single test reaching the level giving a 2% false-positive rate. These failures are heavily weighted towards false positives, with a false positive rate of 20.5%, and a late + missed warning rate of 1%. When the threshold is increased to 6 points, which only occurs when at least two tests register points, the false positive rate drops to 2.5%, at the expense of an increase in the percentage of late and missed disruptions to 4.8%.

The results are somewhat better for the six-level test in figure 7(b). For a warning threshold of 5 points, requiring as per table A2 that only a single test reach the threshold level associated with a 0.5% false-positive rate, the late + missed

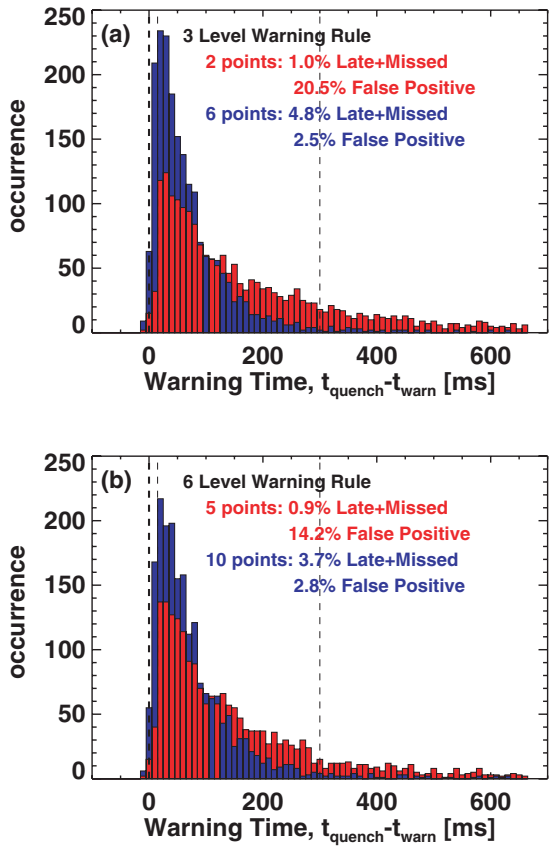

Figure 7. Histograms of warning times. (a) shows the statistical results of the three-level tests, while (b) shows the results of the six-level test.

warning fraction is approximately 1%, while the false-positive rate is still high at 14%. However, increasing the warning level to 10 points, which only occurs when more than a single test registers the crossing of a threshold, leads to a late + missed

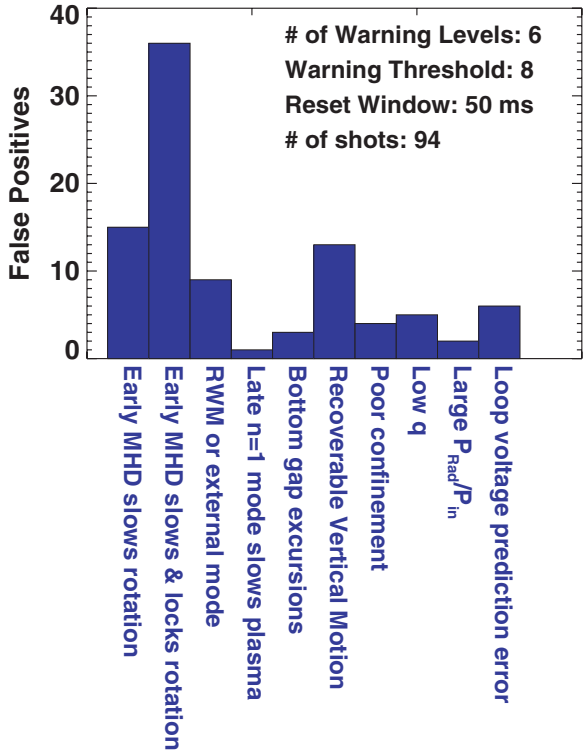

Figure 8. Histogram of false positive causes for the disruption detection algorithm described in sections 4.1–4.3.

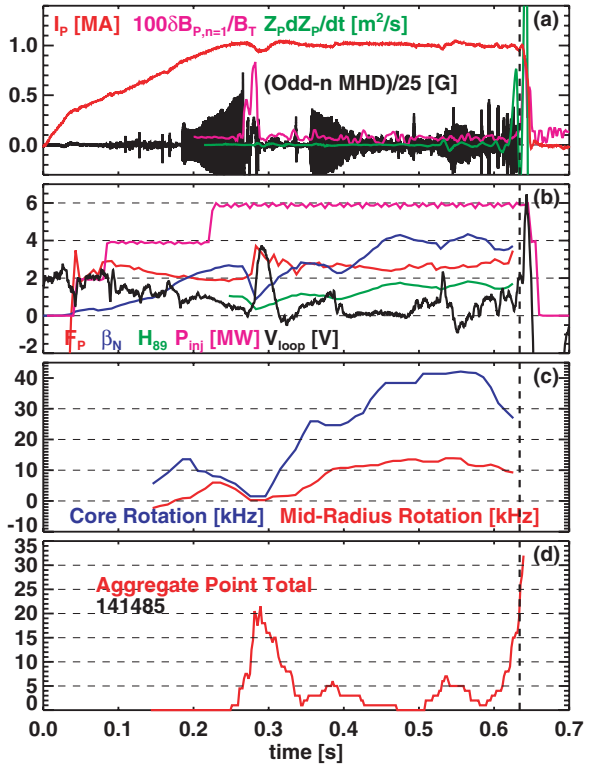

Figure 9. Example waveform for a discharge where an early locked-mode results in a false-positive disruption warning. See figure 6 for a description of the individual panels.

warning rate of 3.7%, with a false-positive rate of only 2.8%. For this last case, the total failure rate is 6.5%.

4.4. Conditions leading to false positives or failed detection

Having established in section 3 that there are many diagnostic indicators that can indicate a disruption is imminent, and in sections 4.1–4.3 that these threshold tests can be combined to form a disruption warning algorithm, it is interesting to consider the conditions under which these tests fail. This will be explored in the present section, concentrating first on the sources of false positives, and then on the late and missed alarms. Here, the false positives and early warnings are defined using the six-level warning rule with a threshold of 8 points to declare a disruption warning; this is intermediate between the cases in figure 7(*b*) and provides a false-positive rate of 4.6%, with a late-warning rate of 2.3%.

Figure 8 provides a histogram explaining the sources of false positives in NSTX. The first two sources provide the largest number of false positives, and are related to the early rotating $n = 1$ MHD modes noted in section 6 of [117], and in figures 6(*a*)–(*d*) of this paper. As shown in that latter figure, these modes often lock to the wall, leading to a disruption. However, we have observed that on occasion, the rotation frequency of those modes will sweep down to nearly zero, but a lock will not occur; the plasma eventually spins-up when the mode dissipates, and the discharge continues. In this case, the false positive is generated by the extreme rotation damping, loss of differential rotation, and observed large-amplitude, low-frequency MHD mode.

There are also cases where the early $n = 1$ mode does indeed lock, but disruption is avoided; an example of this false-positive type is shown in figure 9. The locking itself is indicated by the sudden drop in amplitude of the odd-n rotating MHD signal at $t = 0.27$, followed by the rapid growth of the quasi-stationary $n = 1$ B_P perturbation ($\delta B_{P,n=1}/B_T$ in this figure). Both the core and edge rotation go to zero at $t = 0.28$ s in this example. Furthermore, there is a significant drop in β_N and H_{89}, and large increases in the loop voltage V_{loop} and pressure peaking F_P. Together, these changes result in a large aggregate point total in figure 9(*d*), exceeding the threshold required to declare a warning. However, the discharge recovers from this large event, and the warning is thus classified as a false positive. Note that this discharge does indeed disrupt at later time, and the point total does indeed increase rapidly in the approach to that event.

Returning to figure 8, there are additional causes of false positives. In a few cases, a large quasi-stationary $n = 1$ B_P perturbation grows up, but subsequently vanishes, and the plasma continues on without disruption. There are also examples where large vertical motion is observed, at a level that would generally by disruptive and that thus triggers a disruption warning, but the vertical control system is able to regain control of the equilibrium. There are a few discharges where anomalously poor confinement triggers an alarm, despite the fact that the discharge appears to be stable. There are a few discharges with quite low edge safety factor that are maintained for longer than the 300 ms required to declare a false positive. As noted in section 3.3, the operational low-q limit is set by a host of operational and MHD issues; rare cases where these are avoided and the discharge sustained results in a false positive for the present disruption warning scheme. There are a number of cases where a ramping radiated power is observed, achieving levels that rapidly increase the aggregate point total, but with disruption delayed beyond the 300 ms used to define a false

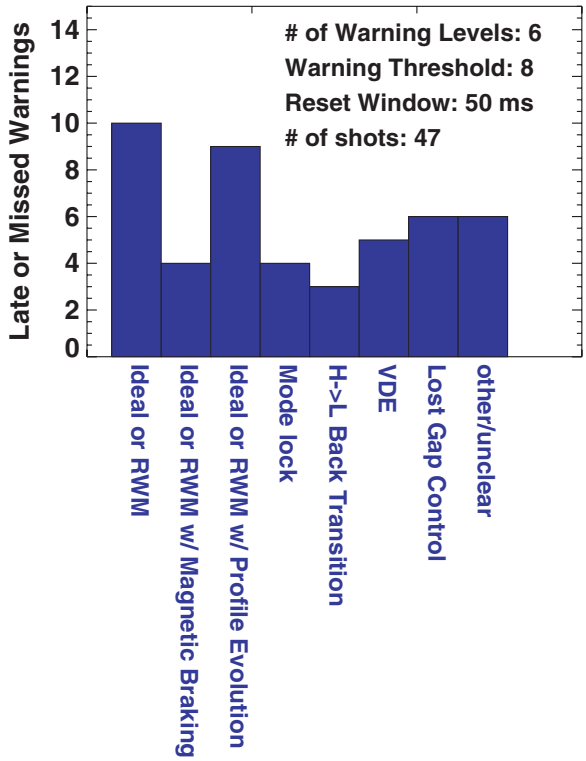

Figure 10. Histogram of causes of late or missed warnings for the disruption detection algorithm described in sections 4.1–4.3.

positive. Finally, there are a few cases where the application of strong heating power early in a lower current discharge results in a strong over prediction of the non-inductive fraction in the model described in section 3.2, and thus under prediction of the loop voltage. This results in a strong contribution to the aggregate point total from the loop voltage ratio test. This could be avoided by a more sophisticated current drive calculation. We note that in most of the cases in figure 8, the plasma is displaying strong signs of distress, and so these 'false positives' may in some cases be better considered as 'disruptions avoided'. Indeed, it appears that in larger plasmas with considerably higher levels of stored energy, the energy lost during many of these events would be sufficient to release large quantities of impurities from the walls, likely resulting in disruption.

The typical causes of disruption for late or missed warnings are shown in figure 10. The first three bars correspond to cases with rapidly growing $n = 1$ MHD modes; these may be RWMs, or more internal modes. In some cases, these modes occur without any clear evolution of the plasma leading to the mode onset. In other cases, either magnetic braking was applied to the plasma, or there was a clear evolution of the pressure peaking to higher values, which may be responsible for the mode onset (note that the level of F_P achieved was not sufficient by itself to trigger a disruption warning). This F_P evolution is often related to undesirable divertor interactions. These cases related to ideal MHD activity make up the largest group of the late warnings.

There are other causes of late warnings indicated in figure 10. The locking of previously rotating MHD modes causes disruptions which are generally quite detectable in advance. However, there were a limited number of cases where the mode locking and subsequent disruption were too fast to detect within the 10 ms time-window set for defining late warnings in this report. H- to L-mode back transitions are typically disruptive in NSTX, due to the higher pressure peaking in L-mode. A limited number of these back transitions and subsequent disruptions occurred too fast for detection. Similarly, there were a small number of VDEs that occurred too quickly for detection given the thresholds used in this example, as well as cases where loss of gap-control resulted in immediate disruption.

There are a few caveats to be considered with figure 10. First, the exact cause of a disruption is not always clear from the available data or physics operator log-book entries. Sometimes, no clear cause is apparent, while in others, multiple causes may be considered. For instance, there are cases where a previously rotating mode locks at the same time that the bottom gap goes to zero, and it is not clear in which bin in figure 10 a case like this should be placed. See [150] for a more thorough discussion of the complicated nature of disruption causes. Also, many of these disruptions would have been detected if the algorithm had been tuned for a smaller late-warning rate, i.e. a lower aggregate point total required to declare a disruption. Doing so, however, would have increased the number of false positives in figure 8.

4.5. Stability of the coefficients over time

A key issue with any on-line disruption detection algorithm is its continued efficacy over time, as the tokamak conditions evolve. In order to address this issue, figure 11 presents the threshold levels for various tests as a function of year. In particular, the threshold levels that result in 10%, 5%, 2%, 1% and 0.5% false-positive rates are indicated with different coloured points, with a label on the right of each figure. These thresholds are independently computed for the data in each of the 2008, 2009 and 2010 run campaigns. The dashed lines show the thresholds based on the complete data set including all three years; these are the values used in determining the point assignment for the tests in sections 4.3 and 4.4.

In some cases, the threshold levels corresponding to low false-positive rates changed considerably over the years. In figure 11(a) for instance, the 0.5% false-positive threshold on $\delta B_{P,n=1}/B_T$ increases by more than a factor of two over the three years under consideration. In this particular case, changes to how the sensor data was processed may contribute to the increased threshold (in particular, an 'AC' compensation step was added to the realtime sensor processing, which reduces some magnetic pickup in the sensor data). However, it appears likely that the increasing use of the RWM control system allowed higher external perturbations to be tolerated. Similarly, the 0.5% false positive thresholds for the H_{89} factor in figure 11(c), the fractional I_P deviation in figure 11(e), and value of $Z_P \cdot dZ_P/dt$ in figure 11(f) all changed substantially over the years. These changes are likely due to variations in the NSTX operating regimes, as noted at the end of section 2. Note also that, as indicated in figure 5, threshold levels with very low false-positive rates often have fairly low total trigger counts, and so the statistical variation in determining these values may be significant. The thresholds associated with higher false-positive rates, however, have better statistics, which may account for their being more constant across the years.

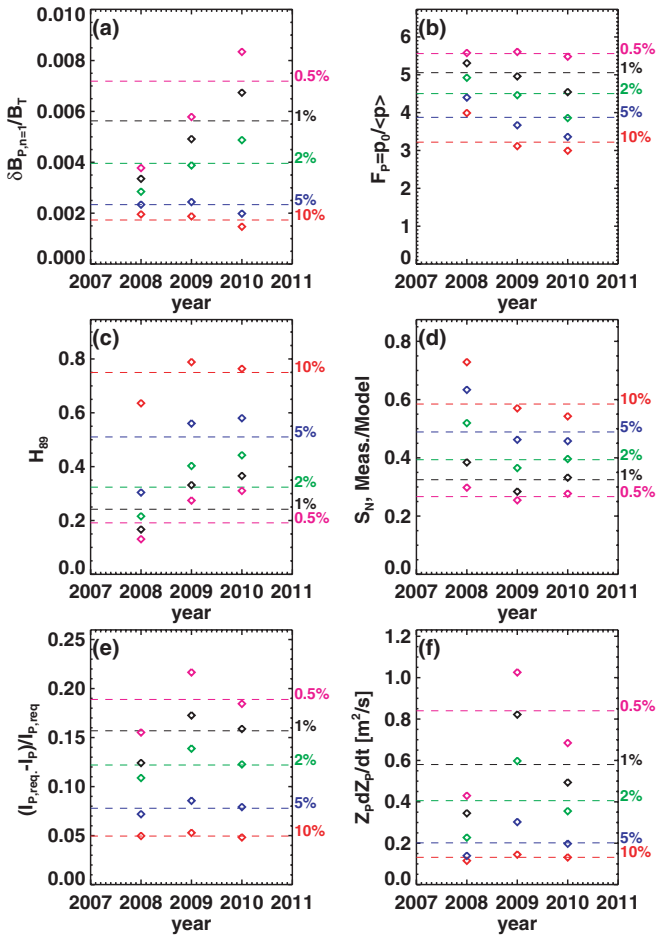

Figure 11. Threshold levels for false-positive rates of 10%, 5%, 2%, 1%, and 0.5%, as a function of year, for six of the tests used in the disruption detection algorithm. The colours denote the false-positive rate, as indicated in the labels on the right of each frame. The tests are (*a*) the normalized quasi-stationary $B_{P,n=1}$ perturbation, (*b*) the pressure peaking, (*c*) normalized global confinement, (*d*) the neutron emission ratio, (*e*) the fractional I_P deviation and (*f*) the quantity $Z_P \cdot dZ_p/dt$.

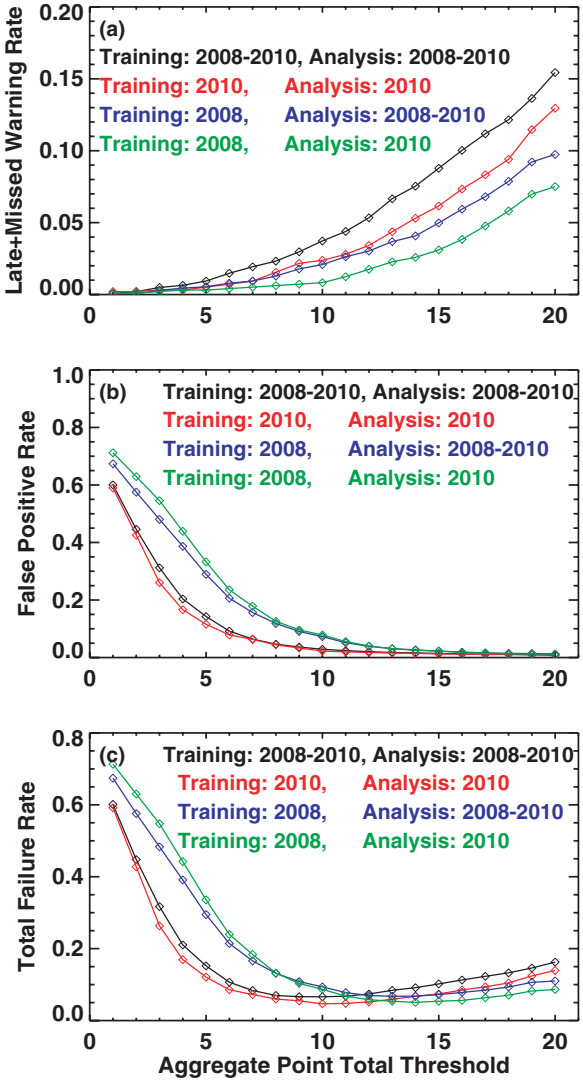

Figure 12. Rates of (*a*) late and missed warnings, (*b*) false positives and (*c*) total failures (sum of false positive and missed or late warnings), as a function of the value of the aggregate point total at which a disruption warning is declared. The different curves correspond to different data sets used to determine the coefficients in the detector or to analyse the failure statistics. See text for additional details, and note the different scales on the figures.

It is worth noting, however, that the threshold levels associated with low false-positive rates can in some cases be fairly constant over the years. This is especially true for the test based on deviations in the neutron rate in figure 11(*d*), and for the test based on thresholds on the pressure peaking factor in figure 11(*b*).

Given the results in figure 11, it is interesting to consider how the disruption detector behaves using different 'training' data sets, that is, different sets of discharges used to determine the point thresholds for the given false-positive rates. The results of this study are indicated in figure 12, showing the rates of late+missed warning, false positive, and their sum (the total failure rate) as a function of the value in the aggregate point total at which a disruption warning is declared. This is for the six-level test with a 50 ms reset window.

The black curves in figure 12 show the cases where the complete three-year data set was both used to determine the coefficients and to evaluate the rates of failure; this is the same data usage as in sections 4.2 and 4.3. As expected from previous discussion, if a disruption warning is declared at fairly low values the aggregate point total, then the late-warning rate can be low, but the false-positive rate may be unacceptably high. Increasing the threshold to 9–10 points results in a minimum total failure rate. Note that this is the value chosen in figure 7(*b*), which displayed reasonably good performance. Increasing the threshold beyond this value results in a gradual increase in the failure rate, as the late-warning rate increases.

The red lines in the figure are based on analysing the 2010 data only, for both determining the coefficients and analysing the various failure rates. There is a reduction in the total failure rate in this case compared to the black curve, presumably due to the smaller variation in the range of operating conditions in the smaller data set.

The other two curves in the figure correspond to cases where only the 2008 data was used to determine the coefficients in the detector: the blue curve uses the entire three-year data set to determines the failure rates, while the green curve is based on the 2010 discharges only in determining the failure rates. Interestingly, there tends to be a slight reduction in the late-warning rate in these cases, apparently because the

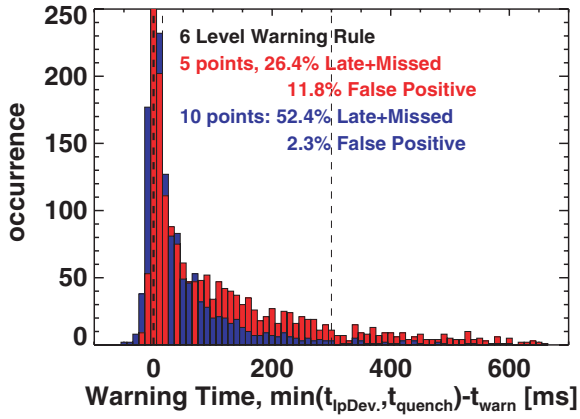

Figure 13. Disruption warning statistics, where the disruption time is defined $t_{dis} = \min(t_{quench}, t_{IpDev})$. See text for further details.

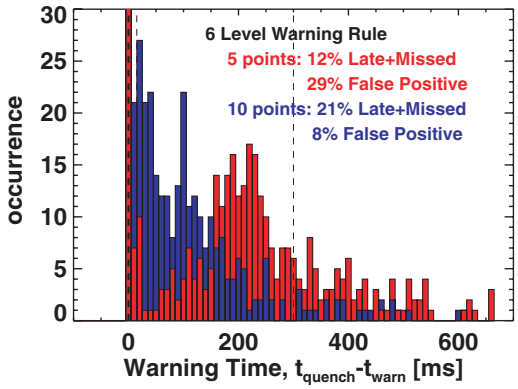

Figure 14. Histogram of warning times, for shots that disrupt during the I_P rampdown.

detector based on 2008 data tended to have more sensitive thresholds. The corollary to the reduced late-warning rate is a large increase in the false-positive rate, especially at lower thresholds in the aggregate point total threshold. However, at larger values of the aggregate point total warning threshold, the performance of the detector is fairly constant, regardless of the data set used for training the detector or analysing the failure rates. In particular, the optimal value of the aggregate point total to use may shift, but the minimum value of the total failure rate is fairly constant.

4.6. Detectability of the disruption-initiating event

The disruption has been defined as the time of the current quench in sections 3 and 4.1–4.4. However, as noted in [117] and in the discussion of figure 1 above, there is often a significant time delay between the instability/event starting the sequence that results in the disruption, and the actual thermal and current quenches. Furthermore, the events during this period provide the basis for many of the disruption signatures used for detection above. Hence, it is interesting to consider how the detection statistics in figure 7 would look if the disruption time is defined as the minimum of either the time of the large negative first I_P deviation, or of the current quench initiation: $t_{dis} = \min(t_{quench}, t_{IpDev})$. For instance, this would be approximately time t_2 in figure 1, or the time indicated by the blue vertical lines in the left and right columns of figure 6 (also see figure 1 of [117] for additional examples).

The results of such a study are shown in figure 13, for the same set of discharges and six-level warning algorithm as in figure 7(*b*). In this case, the number of false positives has decreased slightly, as the event defining the disruption is moved earlier in the discharge while the evolution of the aggregate point total is fixed. More importantly, however, the number of late warnings has increased dramatically, increasing to half of all disruptions in the case with 10 points required for declaring the warning. This can be seen by the increase in the area to the left side of +10 ms in the figure. From this, we conclude that it can be much harder to predict the disruption-initiating event, compared to predicting the thermal/current quench. The implications of this will be discussed in section 5.

4.7. Application of the detection algorithm to disruptions during the i_p rampdown

The analysis in sections 4.3–4.6 presents extensive analysis of the disruption prediction algorithm, from a database of discharges that exhibit disruptions during the I_P flat-top. One additional question that may be asked is with regard to the prevalence of disruption warnings being declared in discharges that do not have a thermal/current quench before the I_P rampdown is initiated. This question is discussed in this section.

As shown in [117], the disruption rate in NSTX was quite high, with ~70% of discharges disrupting at some point during the flat-top and ~25% of discharges disrupting during the current rampdown; a negligible fraction of discharges were ramped down below $I_P = 300$ kA without some large disruptive event. During NSTX operations, the rampdown was on occasion carefully programmed to reduce the stored energy before the plasma current. However, it was more common that the rampdown was initiated by the rapid return to zero of the solenoid current after that coil reached its current limit; this flux swing is in the direction to ramp-down the plasma current, though not necessarily in a manner leading to a smooth plasma current evolution.

The disruption detection algorithm has been applied to a set of 365 neutral beam-heated discharges that disrupt after the I_P rampdown has been initiated, and the results are shown in figure 14. The performance of these algorithms is reasonably good, with ~70% of the disruptions predicted within the $t_{quench} - 10\,\text{ms} > t_{warn} > t_{quench} - 300\,\text{ms}$ window used earlier in this paper for the six-level rule, for a warning declared when the aggregate point total exceeds 10 points.

Compared to the study of flat-top disruption detection in section 4 (see figure 7), the instance of late or missed warnings is increased. The cause of these late warnings can be seen in the left-hand column of figure 15. In this case, the discharge is operating stably until the rampdown is initiated just after $t = 0.85$ (see reversal of V_{loop} after that time in figure 15(*b*)). The discharge shows no signs of distress up to this time, with the MHD signatures in figure 15(*a*) not displaying any modes, the confinement signatures in figure 15(*b*) holding stable, and the rotation maintained in figure 15(*c*); the aggregate point total in figure 15(*d*) does not show any large level before the rampdown is initiated.

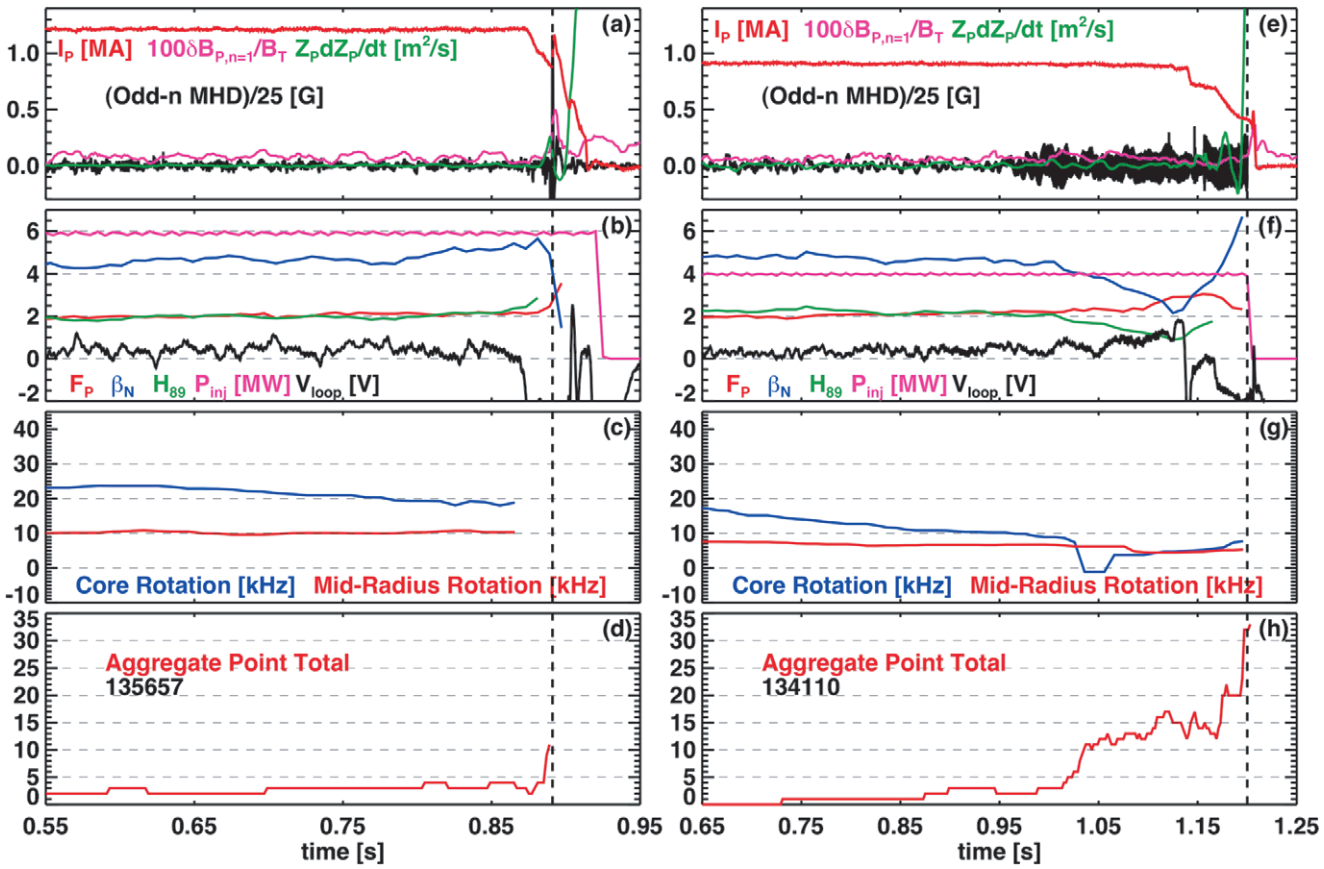

Figure 15. Examples of application of the algorithm to rampdown disruptions. (*a*)–(*d*) show an example with a very short warning time, while (*e*)–(*f*) show an example where the warning would be declared before the rampdown is initiated. See figure 6 for a description of the individual panels, and the text for additional details.

The disruption triggered by the ramp-down occurs at $t = 0.89$, with minimal precursor activity; the aggregate point total only begins to increase a few ms before the I_P spike occurs. As described in section 9 of [117], disruptions initiated by a reversal of the loop voltage accounted for essentially all of the high stored energy disruptions in NSTX, due to the lack of stored energy loss or confinement degradation leading up to the disruption; this lack of precursor activities makes these cases hard to detect in advance. This example shows that the rampdown must be carefully tailored to avoid disruptions, and that the disruption detection algorithm may need to be optimized differently for this phase of the discharge.

The time required to ramp-down the current in typical NSTX plasmas was ∼70–120 ms. This implies that all of the disruption warnings longer than this time in figure 14 actually occur during the I_P flat-top, even though the disruption occurred after the rampdown was initiated. These cases correspond to discharges where the plasma was showing signs of distress indicative of a disruption being imminent, but the rampdown was initiated before the disruption occurred.

An example of this type is shown in figures 15(*e*)–(*h*). In this case, a large $n = 1$ rotating MHD mode is struck at $t = 0.95$, as illustrated in figure 15(*e*). This mode results in a reduction of confinement and an increase in pressure peaking (figure 15(*f*)), and a damping of the rotation with a reduction in rotation shear (figure 15(*g*)). These changes to the plasma result in the aggregate point total in figure 15(*h*) exceeding 10 by $t = 1.05$ s; however, the plasma does not disrupt before the rampdown is initiated at $t = 1.15$. The plasma then disrupts at $t = 1.2$, when that rotating $n = 1$ mode locks to the wall.

Approximately 50% of the cases in this database of rampdown disruptions show the disruption warning occurring before the rampdown is initiated, due to the plasma displaying significant signs of distress before the rampdown as in the right-hand column of figure 15. This and similar examples may be viewed as instances where the warning is triggered but there is no disruption. However, it is likely equally valid to view these as cases where the disruption is preempted by the initiation of the rampdown process, and follows shortly thereafter.

5. Summary and discussion

This paper has given a comprehensive examination of disruption precursors in high-β NSTX plasmas. Key findings include the following.

- There are many diagnostic signals that can provide a basis for the detection of high-β disruptions. These include direct instability detection (vertical motion, stationary and rotating $n = 1$ perturbations), equilibrium characteristics (profile shapes, boundary-wall gaps), and transport indicators (plasma rotation, energy and particle confinement, neutron production, required loop voltage). See sections 3.1–3.3 for more details.
- No single one of these sensors can predict disruptions with sufficiently high fidelity to be used as a stand-alone disruption detector. For any given diagnostic and

threshold level, either the false-positive rate will be too high (in order to minimize late warnings), or the late warning+missed alarm rate will be too high (in order to minimize false positives). See section 4.2 for details.

- A simple threshold-based detection algorithm can be used to detect the majority of disruptions in NSTX. In particular, this algorithm is based on summing the results of a series of threshold tests. The coefficients used in the detection algorithm are based on analysis of the NSTX database. A total failure rate, defined as the sum of false positives, late warnings, and missed warnings, can be as low as 6% when applied to a set of ∼2000 disruptions from a three-year period. See sections 4.1–4.4.
- While the algorithm does a good job of predicting the current quench in advance, it is not as successful at predicting the events that initiate the disruption process. See section 4.6.
- When the detector was applied to a database of ∼360 disruptions during the I_P rampdown, it was found that approximately half of these had the disruption warning declared before the rampdown was initiated. These were cases where the plasma was showing severe signs of distress, but the eventual disruption was preempted by the initiation of the rampdown. See section 4.7.

Considering the fundamental question described at the end of section 1, the answer appears to be that, in general, disruptions in NSTX are detectable in advance. Algorithms that have an extremely low rate of missed or late warnings, however, will often have an unacceptable rate of false positives. This underscores the need to incorporate additional physics and engineering data into the detection scheme, as described later in this section. Note also that while the disruption dynamics in a spherical torus have many similarities to the dynamics in conventional aspect ratio tokamaks, further work is required to demonstrate the applicability of these disruption detection techniques to those larger-A devices.

Looking towards larger next-step devices, there are reasons to believe that the efficiency of these disruption detection schemes may be either better or worse. On the one hand, those devices will typically have only a few operating modes, with limited variation in the time-trajectories of quantities such as the edge-q, boundary shape, β, or the non-inductive current drive sources. This is to be contrasted with the data set used in the present NSTX study, which contains a very large range of elongations, triangularities, edge-q and β values, injected power and non-axisymmetric field configurations. Furthermore, many disruptive actions taken in NSTX to facilitate the research programme would likely result in a controlled termination of the discharge in a larger device, before a disruption occurs; examples include rapid reductions of the heating power or application of large $n = 1$ fields. Finally, the characteristic time-scale of instabilities will be significantly longer than in NSTX, due to the hotter plasma and larger size, providing more absolute time for detection. However, the diagnostic access in those devices will likely be more limited, due to the desire to avoid neutrons streaming through penetrations in the blanket modules and the associated loss in tritium breeding. This reduction in data quality will adversely impact the ability to detect imminent disruptions.

It is interesting to consider some comparisons of this method to neural network based detection. Both methods require that a database of discharges exist, in order to develop the correct coefficients. In the present case, where false-positive rates down to 0.2% are utilized, this implies that at least $1/0.002 = 500$ discharges be available for this determination, though a somewhat different combination of points and false-positive percentages could reduce this number considerably. The primary advantage of the present scheme is that the coefficients so derived are directly traceable to physics issues (those discussed in section 3), unlike in a neural network where a machine learning method develops the coefficients. For this reason, it appears likely that this method may be better able to expand to operating regions not previous explored, or be transferred from one tokamak to another more easily; these potential benefits can only be assessed with additional analysis. On the other hand, the neural network may be better able to discern couplings in the data that lead to improved prediction, compared to the present case, where various pieces of data are evaluated independently. Also, while the present system produces only an 'alarm' trigger, neutral networks can be configured to produce estimates of the time until the disruption [86], a piece of information that can prove quite useful in deciding which mitigation strategy to employ.

An interesting finding of this work is that in the second to final bullet above: that the algorithm is less successful at finding the disruption-initiating event, compared to predicting the current quench. This implies a number of avenues for future research. First, it is clear that better physics is required in the detector in order to identify the causes of imminent disruptions. Potential means of improving the physics fidelity of prediction include the following.

- Realtime $n = 0$ stability assessment could be conducted. For instance, a quantity like ΔZ_{max} [52], indicative of the maximum controllable vertical displacement, could be calculated based on equilibrium quantities determined in realtime. These could then be compared to measurements or estimates of the disturbance spectrum, in order to determine when vertical control has become marginal and the likelihood of a VDE has significantly increased. These calculations could be done against not only the present state of the plasma, but also the projected future state based on realtime current diffusion [151] and equilibrium [37] calculations. If calculations of $n = 1$ stability ever achieve the same level of maturity as those for $n = 0$, then they could also be evaluated in realtime, provided the realtime equilibria are of sufficient quality. Should this not be the case, then a second method can be used for the evaluation of $n = 1$ stability, as described below.
- Resonant field amplification [152] could be used for realtime $n = 1$ stability assessments [153], in order to reduce the likelihood of unanticipated RWM disruptions. These measurements have the advantage of being intrinsically sensitive to the details of kinetic stabilization physics [143, 154], effects that are only recently being considered in sophisticated calculations [139–142, 155] not presently feasible in a realtime manner. Note that realtime RFA measurements have been used for feedback control of the neutral beam power in DIII-D, in order to maintain constant proximity to the $n = 1$ stability

Table A1. Thresholds used in the three-level warning scheme, as well as a short description of each test.

Test	1 pt 5% FPR	2 pt 2% FPR	3 pt 0.5% FPR	Description
$100\delta B_{P,n=1}/B_T >$	0.23	0.40	0.72	Quasi-static $n = 1$ B_P perturbation too large
$\lvert dZ_P/dt\rvert$ (m s^{-1}) >	3.93	6.54	9.01	Absolute value of vertical velocity too large
$F_P >$	3.88	4.50	5.56	MHD pressure peaking too high
$l_i >$	0.84	0.90	1.02	Internal inductance too high
$H_{89} <$	0.51	0.32	0.19	Normalized confinement too low
V_{loop}, meas./model >	13.50	24.00	*	Ratio of measured to modelled loop voltage too large
S_N, meas./model <	0.49	0.39	0.27	Ratio of measured to modelled neutron emission too small
Bottom Gap (m) <	0.03	*	*	Gap between plasma and divertor floor too small
$(I_{P,req.} - I_P)/I_{P,req} >$	0.08	0.12	0.19	Fraction I_P error too large
$q_0 <$	0.87		*	Central safety factor too low
$q* <$	2.82	2.67	2.51	Cylindrical safety factor too low
$q_{95} <$	7.05	6.50	6.02	Edge safety factor (q_{95}) too low
$100F_{T,core}/F_A <$	1.45	0.39	*	Normalized core rotation frequency too low
$F_{T,core} - F_{T,mid\text{-}radius}$ (Hz) <	1423	*	*	Core to mid-radius differential rotation too low
dn_{line}/dt (10^{16} cm^{-2} s^{-1}) <	−2.58	−5.58	−9.95	Line-density drop too large
$P_{rad}/P_{heat} >$	0.25	0.31	0.40	Radiated power too large compared to heating power
$Z_P \cdot dZ_P/dt >$	0.20	0.41	0.84	Product of vertical position and vertical velocity too large
Coil fault >	0.10	0.20	0.30	Power supply control software declares a fault

Table A2. Thresholds used in the six-level warning scheme.

Test	1 pt 10% FPR	2 pt 5% FPR	3 pt 2% FPR	4 pt 1% FPR	5 pt 0.5% FPR	7 pt 0.2% FPR
$100\delta B_{P,n=1}/B_T >$	0.17	0.23	0.40	0.56	0.72	0.88
$\lvert dZ_P/dt\rvert$ (m s^{-1}) >	2.88	3.93	6.54	7.92	9.01	9.70
$F_P >$	3.22	3.88	4.50	5.06	5.56	6.49
$l_i >$	0.77	0.84	0.90	0.97	1.02	1.14
$H_{89} <$	0.75	0.51	0.32	0.24	0.19	0.13
V_{loop}, meas./model >	8.90	13.50	24.50	43.50	*	*
S_N, meas./model <	0.58	0.49	0.39	0.32	0.27	0.19
Bottom Gap (m) <	0.06	0.03	*	*	*	*
$(I_{P,req.} - I_P)/I_{P,req} >$	0.05	0.08	0.12	0.16	0.19	0.26
$q_0 <$	1.01	0.87	*	*	*	*
$q* <$	2.97	2.82	2.67	2.56	2.51	2.38
$q_{95} <$	7.52	7.05	6.50	6.18	6.02	5.69
$100F_{T,core}/F_A <$	2.21	1.45	0.39	*	*	*
$F_{T,core} - F_{T,mid\text{-}radius}$ (kHz) <	2161	1423	*	*	*	*
dn_{line}/dt (10^{16} cm^{-2} s^{-1}) <	−1.34	−2.58	−5.58	−8.27	−9.95	−13.22
$P_{rad}/P_{heat} >$	0.20	0.25	0.31	0.35	0.40	0.47
$Z_P \cdot dZ_P/dt >$	0.13	0.20	0.41	0.58	0.84	1.17
Coil fault >	0.10	0.20	0.30	0.40	0.50	0.60

limit [156], indicating the applicability of the technique to realtime control.

- Ultimately, many of these tests should be framed in terms of loss of control by critical actuators. In the present case, the plasma current, plasma vertical position, and the $n = 1$ distortion are under feedback control. Three of the tests in section 3 essentially quantify the extent to which the actuators are no-longer capable of controlling these quantities: the thresholds on $(I_{P,req} - I_P)/I_P$ represent current deviations too large to correct, the thresholds on $Z_P \cdot dZ_P/dt$ represent vertical motion too well developed to control, and the thresholds on $\delta B_{P,n=1}$ represent $n = 1$ distortions too large to correct. However, the thresholds used here are based on experimental measurements, and not on models of control. Ideally, models of the plasma, feedback system, and actuator capabilities would be used to provide more first-principle loss of control thresholds, the crossing of which would mandate a significant change to the discharge trajectory or complete discharge termination.

Beyond these considerations, the reduced ability to detect the disruption-initiating event leads to questions about how disruption detection may differ in devices which do not have any capability for solenoid induction. In all experiments in NSTX, the solenoid applied a considerable voltage during the pre-disruption phase, due to the loss of non-inductive current drive. This is an effective means of driving current and supplying power to the relatively cold plasma, and can result in a substantial extension of the disruption phase, during which disruption detection can often be easily accomplished. However, if there were no solenoid, as is common in the design of ST configurations for the FNSF/CTF [157–160], pilot plant

[161], or reactor [162, 163] missions, the disruption process may be much faster, and detection significantly more difficult (see [164] for a description of the CTF/FNSF mission). The extensive non-inductive capabilities [165] of NSTX-Upgrade [128] should allow these studies of disruption detection in high-β, 100% non-inductive fraction discharges with solenoid-based I_P feedback control disabled.

Acknowledgments

This work was funded by the United States Department of Energy under contract DE AC02 09CH11466. The authors would like to thank Bill Heidbrink for help with the 0D slowing down model used for predicting the neutron emission.

Appendix. Parameters of the disruption detection algorithm

Tables A1 and A2 provide the complete point-level assignments for the three- and six-level tests described in section 4.3. These coefficients are determined by analysis of the full 2026 discharge data set from the 2008–2010 campaigns. The first 17 rows show the thresholds for tests based on physics data. The last row shows a quantity called 'coil fault'. This value is compared against the output of the power supply control software, which is either a 0 (good) or a 1 (fault). Hence, if a fault is declared, the largest possible point value will be added to the aggregate point total. The final column of table A1 contains a short description of the tests.

References

[1] Wesson J. 1997 *Tokamaks* (Oxford: Clarendon)
[2] Peng Y.K.M. and Strickler D.J. 1986 *Nucl. Fusion* **26** 769
[3] ITER Physics Experts Groups 1999 *Nucl. Fusion* **39** 2251
[4] Hender T.C. *et al* 2007 Progress in the ITER Physics Basis: Chapter 3. MHD stability, operational limits and disruptions 2007 *Nucl. Fusion* **47** S128–202
[5] Riccardo V. *et al* 2002 *Plasma Phys. Control. Fusion* **44** 905
[6] Riccardo V., Loarte A. and JET EFDA Contributors 2005 *Nucl. Fusion* **45** 1427
[7] Arnoux G. *et al* 2008 *Nucl. Fusion* **49** 085038
[8] Riccardo V. *et al* 2010 *Plasma Phys. Control. Fusion* **52** 024018
[9] Federici G. 2006 *Phys. Scr.* **T124** 1
[10] Loarte A. *et al* 2007 *Phys. Scr.* **T128** 222
[11] Humphreys D.A. and Whyte D.G. 2000 *Phys. Plasmas* **7** 4057
[12] Riccardo V., Barabaschi P. and Sugihara M. 2005 *Plasma Phys. Control. Fusion* **47** 117
[13] Wesley J.C. *et al Proc. 21st Int. Conf. on Fusion Energy 2006 (Chengdu, People's Republic of China 2006)* (Vienna: IAEA) CD-ROM file IT/P1-21 and www-naweb.iaea.org/napc/physics/FEC/FEC2006/html/index.htm
[14] Gerhardt S.P., Menard J.E. and the NSTX Team 2009 *Nucl. Fusion* **49** 025005
[15] Wesley J.C. *et al* 2010 *Proc. 23rd Int. Conf. on Fusion Energy (Daejeon, South Korea 2010)* (Vienna: IAEA) CD-ROM file ITR/P1-26 and www-naweb.iaea.org/napc/physics/FEC/FEC2010/html/index.htm
[16] Shibata Y. *et al* 2010 *Nucl. Fusion* **50** 025012
[17] Sugihara M. *et al* 2007 *Nucl. Fusion* **47** 337
[18] Rosenbluth M.N. and Putvinski S.V. 1997 *Nucl. Fusion* **37** 1355
[19] Yoshino R., Tokuda S. and Kawano Y. 1999 *Nucl. Fusion* **39** 151
[20] Yoshino R. and Tokuda S. 2000 *Nucl. Fusion* **40** 1293
[21] Gill R.D. *et al* 2002 *Nucl. Fusion* **42** 1039
[22] Tamai H. *et al* 2002 *Nucl. Fusion* **42** 290
[23] Forster M. *et al* 2011 *Nucl. Fusion* **51** 043003
[24] Hollmann E.M. *et al* 2011 *Nucl. Fusion* **51** 103026
[25] Eidietis N.W. *et al* 2012 *Phys. Plasmas* **19** 056109
[26] James A.N. *et al* 2012 *Nucl. Fusion* **52** 013007
[27] Sizyuk V. and Hassanein A. 2009 *Nucl. Fusion* **49** 095003
[28] Strait E.J. *et al* 1991 *Nucl. Fusion* **31** 527
[29] Granetz R.S. *et al* 1996 *Nucl. Fusion* **36** 545
[30] Neyatani Y. *et al* 1999 *Nucl. Fusion* **39** 559
[31] Counsell G.F. *et al* 2007 *Plasma Phys. Control. Fusion* **49** 435
[32] Riccardo V. *et al* 2009 *Nucl. Fusion* **49** 055012
[33] Pautasso G. *et al* 2011 *Nucl. Fusion* **51** 043010
[34] Eidietis N.W. *et al* 2011 *Nucl. Fusion* **51** 073034
[35] Gerhardt S.P. *et al* 2012 *Nucl. Fusion* **52** 063005
[36] Gerhardt S.P. 2013 *Nucl. Fusion* **53** 023005
[37] Ferron J.R. *et al* 1998 *Nucl. Fusion* **38** 1055
[38] Gates D.A. *et al* 2006 *Nucl. Fusion* **46** 17
[39] Zornig N.H. *et al* 1996 Experimental results using the JET real time power control system *Proc. 19th Symp. on Fusion Technology* (*Lisbon, Portugal, 16–20 September 1996*, vol 1) pp 705–8
[40] Scoville T. *et al* 2007 *Fusion Eng. Des.* **82** 1045
[41] Gerhardt S.P. *et al* 2012 *Fusion Sci. Technol.* **61** 11
[42] Moreau D. *et al* 2003 *Nucl. Fusion* **43** 870
[43] Joffrin E. *et al* 2007 *Nucl. Fusion* **47** 1664
[44] Moreau D. *et al* 2008 *Nucl. Fusion* **48** 106001
[45] Moreau D. *et al* 2011 *Nucl. Fusion* **51** 063009
[46] Garofalo A.M., LaHaye R.J. and Scoville J.T. 2002 *Nucl. Fusion* **42** 1335
[47] Menard J.E. *et al* 2010 *Nucl. Fusion* **50** 045008
[48] Lazarus E.A. *et al* 1991 *Phys. Fluids* **3** 2220
[49] Hofmann F. *et al* 1997 *Nucl. Fusion* **37** 681
[50] Hofmann F. *et al* 2000 *Nucl. Fusion* **40** 767
[51] Kessel C.E., Heitzenroeder P. and Jun C. 2001 *Nucl. Fusion* **41** 953
[52] Humphreys D.A. *et al* 2009 *Nucl. Fusion* **49** 115003
[53] Bondeson A. and Ward D.J. 1994 *Phys. Rev. Lett.* **72** 2709
[54] Ward D.J. and Bondeson A. 1995 *Phys. Plasmas* **2** 1570
[55] Strait E.J. *et al* 2004 *Phys. Plasmas* **11** 2505
[56] Sabbagh S.A. *et al* 2006 *Phys. Rev. Lett.* **97** 045004
[57] Garofalo A.M. *et al* 2007 *Nucl. Fusion* **47** 1121
[58] Okabayashi M. *et al* 2005 *Nucl. Fusion* **45** 1715
[59] Okabayashi M. *et al* 2009 *Nucl. Fusion* **49** 125003
[60] Liu Y. *et al* 2009 *Phys. Plasmas* **16** 056113
[61] Sauter O. *et al* 1997 *Phys. Plasmas* **4** 1654
[62] La Haye R.J. *et al* 2006 *Phys. Plasmas* **13** 055501
[63] Isayama A. *et al* 2000 *Plasma Phys. Control. Fusion* **42** L37
[64] Petty C.C. *et al* 2004 *Nucl. Fusion* **44** 243
[65] Maraschek M. *et al* 2005 *Nucl. Fusion* **45** 1369
[66] Humphreys D.A. *et al* 2006 *Phys. Plasmas* **13** 056113
[67] Volpe F. *et al* 2009 *Phys. Plasmas* **16** 102502
[68] Esposito B. *et al* 2011 *Plasma Phys. Control. Fusion* **53** 124035
[69] Esposito B. *et al* 2009 *Nucl. Fusion* **49** 065014
[70] Esposito B. *et al* 2011 *Nucl. Fusion* **51** 083051
[71] Whyte D.G. *et al* 2002 *Phys. Rev. Lett.* **89** 055001
[72] Whyte D.G. *et al* 2003 *J. Nucl. Mater.* **313–316** 1239
[73] Granetz R.S. *et al* 2007 *Nucl. Fusion* **47** 1086
[74] Hollmann E.M. *et al* 2008 *Nucl. Fusion* **49** 115007
[75] Pautasso G. *et al* 2007 *Nucl. Fusion* **47** 900
[76] Pautasso G. *et al* 2009 *Plasma Phys. Control. Fusion* **51** 124056
[77] Reux C. *et al* 2010 *Nucl. Fusion* **50** 095006
[78] Hollmann E.M. *et al* 2010 *Phys. Plasmas* **17** 056117
[79] Bakhtiari M. *et al* 2011 *Nucl. Fusion* **51** 063007
[80] Commaux N. *et al* 2011 *Nucl. Fusion* **51** 103001
[81] Commaux N. *et al* 2010 *Nucl. Fusion* **50** 112001

[82] Combs S.K. *et al* 2009 Massive pellet and rupture disk testing for disruption mitigation applications *36th Int. Conf. on Plasma Science and 23rd Symp. on Fusion Engineering (IEEE Conference Publications)* http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5226512
[83] Saint-Laurent F. *et al* 2012 Overview of runaway electrons control and mitigation experiments on Tore Supra and lessons learned in view of ITER *Proc. 24th Int. Conf. on Fusion Energy (San Diego, CA, 2012)* paper EX/P8-06 www-naweb.iaea.org/napc/physics/FEC/FEC2012/papers/226_EXP806.pdf
[84] de Vries P.C., Johnson M.F., Segui I. and JET EFDA Contributors 2009 *Nucl. Fusion* **49** 055011
[85] Pautasso G. *et al* 2002 *Nucl. Fusion* **42** 100
[86] Windsor C.G. *et al* 2005 *Nucl. Fusion* **45** 337
[87] Cannas B. *et al* 2010 *Nucl. Fusion* **50** 075004
[88] Wroblewski D., Jahns G.L. and Leuer J.A. 1997 *Nucl. Fusion* **37** 725
[89] Sengupta A. and Ranjan P. 2000 *Nucl. Fusion* **40** 1993
[90] Sengupta A. and Ranjan P. 2001 *Nucl. Fusion* **41** 487
[91] Hernandez J.V. *et al* 1996 *Nucl. Fusion* **36** 1009
[92] Vannucci A., Oliveira K.A. and Tajima T. 1999 *Nucl. Fusion* **39** 255
[93] Cannas B. *et al* 2004 *Nucl. Fusion* **44** 68
[94] Cannas B. *et al* 2007 *Nucl. Fusion* **47** 1559
[95] Yoshino R. *et al* 2003 *Nucl. Fusion* **43** 1771
[96] Yoshino R. 2005 *Nucl. Fusion* **45** 1232
[97] Troyon F. *et al* 1984 *Plasma Phys. Control. Fusion* **26** 209
[98] Strait E.J. 1994 *Phys. Plasmas* **1** 1415
[99] Greenwald M. *et al* 1988 *Nucl. Fusion* **28** 2199
[100] Greenwald M. *et al* 2002 *Plasma Phys. Control. Fusion* **44** R27
[101] Ono M. *et al* 2000 *Nucl. Fusion* **40** 557
[102] Stevenson T. *et al* 2002 A neutral beam injector upgrade for NSTX *PPPL Report 3651*
[103] Wilson J.R. *et al* 2003 *Phys. Plasmas* **10** 1733
[104] LeBlanc B.P. *et al* 2003 *Rev. Sci. Instrum.* **74** 1659
[105] Bell R.E. and Feder R. 2010 *Rev. Sci. Instrum.* **81** 10D724
[106] Levinton F.M. *et al* 1990 *Rev. Sci. Instrum.* **61** 2914
[107] Levinton F. and Yuh H. 2008 *Rev. Sci. Instrum.* **79** 10F522
[108] Sabbagh S.A. *et al* 2001 *Nucl. Fusion* **41** 1601
[109] Sabbagh S.A. *et al* 2006 *Nucl. Fusion* **46** 635
[110] Lao L.L. 1985 *Nucl. Fusion* **25** 1611
[111] Gates D.A. *et al* 2004 *Rev. Sci. Instrum.* **75** 5090
[112] Heidbrink W.W. *et al* 2003 *Nucl. Fusion* **43** 883
[113] Sontag A.C. *et al* 2005 *Phys. Plasmas* **12** 056112
[114] Sabbagh S.A. *et al* 2004 *Nucl. Fusion* **44** 560
[115] Sontag A.C. *et al* 2007 *Nucl. Fusion* **47** 1005
[116] Sabbagh S.A. *et al* 2010 *Nucl. Fusion* **50** 025020
[117] Gerhardt S.P. *et al* 2013 *Nucl. Fusion* **53** 043020
[118] Kugel H. *et al* 2008 *Phys. Plasmas* **15** 056118
[119] Bell M.G. *et al* 2009 *Plasma Phys. Control. Fusion* **51** 124054
[120] Kugel H.W. *et al* 2009 *Fusion Eng. Des.* **84** 1125
[121] Gates D.A. *et al* 2006 *Nucl. Fusion* **46** S22
[122] Gates D.A. *et al* 2006 *Phys. Plasmas* **13** 056122
[123] Gerhardt S.P. *et al* 2011 *Nucl. Fusion* **51** 073031
[124] Yushmanov P.N. *et al* 1990 *Nucl. Fusion* **30** 1999
[125] ITER Physics Experts Groups 1999 *Nucl. Fusion* **39** 2175
[126] Fredrickson E.D. *et al* 2006 *Phys. Plasmas* **13** 056109
[127] Gerhardt S.P. *et al* 2011 *Nucl. Fusion* **51** 033004
[128] Menard J.E. *et al* 2012 *Nucl. Fusion* **52** 083015
[129] Bickerton R.J., Connor J.W. and Taylor J.B. 1971 *Nature* **229** 110
[130] Galeev A.A. 1971 *Sov. Phys.—JETP* **32** 752
[131] Zarnstorff M.C. and Prager S.C. 1984 *Phys. Rev. Lett.* **53** 454
[132] Sauter O., Angioni C. and Lin-Liu Y.R. 1999 *Phys. Plasmas* **6** 2834
[133] Peeters A.G. 2000 *Plasma Phys. Control. Fusion* **42** B231
[134] Ohkawa T. 1970 *Nucl. Fusion* **10** 185
[135] Fisch N.J. 1987 *Rev. Mod. Phys.* **59** 175
[136] Lin-Liu Y.R. and Miller R.L. 1995 *Phys. Plasmas* **2** 1666
[137] Hawryluk R.J. *et al* 1980 An empirical approach to tokamak transport *Physics of Plasmas Close to Thermonuclear Conditions* vol 1 ed B. Coppi *et al* (Brussels: CEC) pp 19–46
[138] Pankin A. *et al* 2004 *Comput. Phys. Commun.* **159** 157
[139] Berkery J.W. *et al* 2010 *Phys. Rev. Lett.* **104** 035003
[140] Berkery J.W. *et al* 2010 *Phys. Plasmas* **17** 082504
[141] Berkery J.W. *et al* 2011 *Phys. Rev. Lett.* **106** 075004
[142] Berkery J.W. *et al* 2011 *Phys. Plasmas* **18** 072501
[143] Berkery J.W. *et al* 2012 Global mode control and stabilization for disruption avoidance in high-β ST plasmas *Proc. 24th Int. Conf. on Fusion Energy (San Diego, CA, 2012)* paper EX/P8-07 www-naweb.iaea.org/napc/physics/FEC/FEC2012/index.htm
[144] Gerhardt S.P. *et al* 2010 *Plasma Phys. Control. Fusion* **52** 104003
[145] Menard J.E. *et al* 2005 *Nucl. Fusion* **45** 539
[146] Menard J.E. *et al* 2006 *Phys. Rev. Lett.* **97** 095002
[147] Gerhardt S.P. *et al* 2009 *Nucl. Fusion* **49** 032003
[148] Menard J.E. *et al* 2004 *Phys. Plasmas* **11** 639
[149] Kolemen E. *et al* 2012 Vertical stability of NSTX and NSTX-U *Proc. 24th Int. Conf. on Fusion Energy (San Diego, CA, 2012)* paper EX/P4-28 www-naweb.iaea.org/napc/physics/FEC/FEC2012/index.htm
[150] de Vries P.C. *et al* 2011 *Nucl. Fusion* **51** 053018
[151] Felici F. *et al* 2011 *Nucl. Fusion* **51** 083052
[152] Reimerdes H. *et al* 2006 *Phys. Plasmas* **13** 056107
[153] Reimerdes H. *et al* 2005 *Nucl. Fusion* **45** 368
[154] Reimerdes H. *et al* 2011 *Phys. Rev. Lett.* **106** 215002
[155] Liu Y. *et al* 2008 *Phys. Plasmas* **15** 112503
[156] Hanson J.M. *et al* 2012 *Nucl. Fusion* **52** 013003
[157] Peng Y.-K.M. *et al* 2005 *Plasma Phys. Control. Fusion* **47** B263
[158] Voss G.M. *et al* 2008 *Fusion Eng. Des.* **83** 1648
[159] Peng Y.-K.M. *et al* 2009 *Fusion Sci. Technol.* **56** 957
[160] Stambaugh R.D. *et al* 2010 Candidates for a Fusion Nuclear Science Facility (FDF and ST-CTF) *37th EPS Conf. on Plasma Physics (Dublin, Ireland 2010)* Paper P2.110 http://ocs.ciemat.es/EPS2010PAP/pdf/P2.110.pdf
[161] Menard J.E. *et al* 2011 *Nucl. Fusion* **51** 103014
[162] Najmabadi F. and the ARIES Team 2003 *Fusion Eng. Des.* **65** 143
[163] Wilson H.R. *et al* 2004 *Nucl. Fusion* **44** 917
[164] Abdou M. 1995 *Fusion Eng. Des.* **27** 111
[165] Gerhardt S.P. *et al* 2012 *Nucl. Fusion* **52** 083020