1. Introduction

The discovery of Rossby waves (Rossby, 1939) in the Earth’s atmosphere led to great advances in the ability to forecast our planet’s weather patterns. It is the combination of mean west to east atmospheric flow and (finite amplitude) Rossby waves in the atmosphere that create “jet streams” at midlatitudes. Understanding their interaction and the resulting longitudinal structure allows for accurate prediction of how synoptic weather patterns evolve and propagate to the east. In effect the “jet stream” steers our Earth’s weather from location to location in midlatitudes.
A spectacular recent example (January 2019) of Rossby waves perturbing the jet stream to permit the creation of a “polar vortex” which led to an episode of record-breaking cold temperatures in the American upper midwest (https://weather.com/forecast/national/news/2019-01-28-polar-vortex-midwest-arctic-air-coldest-two-decades). In this case Rossby waves of longitudinal wave numbers of one and two resulted in the axis of the polar vortex moving toward lower latitudes as well as deforming the vortex into a fluted shape. These effects combined to result in unusually strong transport of very cold polar air into the mid-west that caused tropospheric temperatures to plummet. The success of modern assimilative meteorology permitted a prediction of this event a few days ahead. However, even with the warnings that were generated by the forecast there were at least 25 deaths directly attributed to the cold outbreak. If this event had not been so well forecast we may be safe in speculating that the death toll would surely have been higher.

Analogous Rossby wave effects have also been identified as responsible for outbreaks of particularly hot weather as well as floods in several vulnerable areas of the world, including California, western Canada, the desert South-west, Mid-Atlantic states, and across Europe and Asia, particularly Russia and Pakistan (see, e.g., Mann, 2019).

So what is a Rossby wave? Most readers are familiar with other waves found in atmospheres, such as sound and gravity waves. They both have simple restoring forces to create an oscillation: pressure and buoyancy, respectively. Rossby waves originate from Coriolis forces in a rotating spherical shell, but they are somewhat peculiar in that the restoring force is actually perpendicular to the flow, rather than directly opposite to it. To help the reader unfamiliar with Rossby waves, we provide in Figure 1 a simple schematic that illustrates the mechanism of Rossby waves. We give a more detailed discussion of Rossby waves in terms of conservation of vorticity, which is the traditional approach in the geophysical fluid dynamics community, in section 2.

There is accumulating observational evidence that the Sun’s atmosphere is also home to Rossby waves (see, e.g., McIntosh et al., 2017, and references therein; see also Krista et al., 2018; Loptien et al., 2018; Hanasoge & Mandal, 2019; Liang et al., 2019; McIntosh et al., 2019). Compared to the Earth’s waves, solar Rossby waves appear to be strongly influenced by the presence of magnetic fields—those are almost ubiquitous in the Sun’s atmosphere. It has recently been demonstrated that the combination of these magnetized Rossby waves and the Sun’s differential rotation coexisting with strong magnetic fields can simulate surface synoptic magnetic activity patterns and their evolution (Dikpati, Belucz, et al., 2018). Analogous to the Earth’s jet stream, these longitudinal magnetic patterns can provide a physical framework to predict the future state of the Sun’s magnetic field from which space weather can be forecast.

The primary aim of this paper is to consolidate some recent, potentially significant, advances in understanding and estimating the spatiotemporal patterns of solar magnetic flux emergence, and convey the importance of solar Rossby waves. They are likely the origin of such spatiotemporal solar activity variations that produce correspondingly large variations in space weather with the intermediate time scale we have defined here. We present a promising future outlook for progress in understanding the role of Rossby waves in determining space weather and hence, motivate the readers to join in future research in this area.

The text which follows is arranged in the following way: section 2 reviews relevant aspects of space weather, with particular focus on time scales characteristic of Rossby waves in the Sun, followed by a few examples of how understanding Rossby waves in the Earth’s lower atmosphere has led to substantial improvements in forecasting major weather events. This comparison is intended to interest space weather scientists and forecasters in the potential for using Rossby waves to enhance our ability to forecast seasons of solar activity and space weather several months ahead. Section 3 reviews the basics of Rossby waves, which is aimed at space weather readers who may be less familiar with them. Section 4 summarizes what we know from observations currently about Rossby waves in the Sun. Section 5 describes theoretical modeling developments applied to solar Rossby waves, both hydrodynamic and magnetohydrodynamic varieties, including neutral and unstable waves, nonlinear waves and interactions with differential rotation and toroidal fields, and the recently discovered tachocline nonlinear oscillations (TNOs).

All of these theoretical developments provide the fundamental tools needed to build forecasting models that include modern data assimilation, so that surface magnetic observations can be coupled to global tachocline magnetohydrodynamics (MHD) to forecast the state of Rossby waves several months ahead, as well as the surface magnetic patterns that should be largely determined by these waves. These ongoing developments, as well as the utility of forecasting Rossby waves, are discussed in sections 6 and 7.
2. Rossby Waves and Space Weather

The term "space weather" was invented to describe the conditions on the outer edge of the terrestrial system, including the near-Earth space, the Earth's upper atmosphere, the solar wind, and all events on the Sun that can affect the space-borne and ground-based technological system. Society's economic activity and national security as well as human health and safety face serious risks of disruption and damage due to powerful space weather events. All these impacts have been extensively documented in studies by the U.S. National Academy of Sciences (Space Studies Board and National Research Council, 2009) and the UK Royal Engineering Academy (Cannon et al., 2013).

In addition to the continuous emission of plasma streams (solar wind), the Sun sporadically sends out billions of tons of energetic plasma particles and magnetic field in the form of coronal mass ejections (CMEs). Those CMEs traverse interplanetary space rapidly and can interact with Earth's magnetic field with the potential to cause significantly hazardous conditions in our atmosphere. The chain of physical processes...
that transmit the effects of CMEs and flares into the Earth’s atmosphere where they can adversely affect society through disrupting and damaging space-weather-sensitive industries and national security systems is very complex, with many scientific questions still to be answered. Schrijver et al. (2015) has laid out a comprehensive “road map,” commissioned by COSPAR and ILWS, for improving our understanding of these processes, particularly on short time scales, that is, minutes to hours and up to a few days. The reader is referred to this study to get a picture of future scientific priorities on these time scales. Our emphasis in this paper is on the time scale of a few weeks to several months up to a year or so. This starts from the long-end of the range considered in Schrijver et al. (2015) in the preparation of their road map. Activity occurring with “seasonal/subseasonal” time scales and their impact on the space weather was not perceived at the time of the preparation of the COSPAR road map. It is these time scales that show evidence of solar Rossby waves in the Sun. Understanding the origins of bursts of solar activity on these time scales therefore complements and extends future studies called for in Schrijver et al. (2015). At the same time, these time scales are much shorter than that of sunspot cycles, that is, decadal and millennial time scales, which we also do not emphasize here. Theories to explain solar cycle features and their impact on the Earth must consider magnetohydrodynamic dynamo theory; there already exists discussions on the understanding and forecasting of solar activity on these much longer time scales (see, e.g., Usoskin, 2017).

This line of research leads to the realization that global-scale solar magnetic flux emergence is governed by the nonlinear dynamics of solar Rossby waves in the solar “tachocline.” The tachocline is a shear layer, with a thickness of no more than 4% of the solar radius, that straddles the base of the solar convection zone just above the rigidly rotating radiative interior (for summary, see Antia et al., 1998). Observations and simulations of the Sun’s Rossby waves could lead to a measurable increase in predictive capability for extreme solar activity on time scales of weeks to months and up to a year—especially when coupled with assimilation of magnetic field information.

Space weather events occur on a wide range of time scales, including (i) short time (hours-to-days) events, (ii) solar seasonal/subseasonal (6–18 months) episodic bursts of activity, and (iii) decadal/century time-scale phenomena. There is interest in forecasting on all three time scales; however in this paper, we will focus on items (i) and (ii), and will omit discussion on decadal and longer time scales which are extensively covered in the broader scientific literature (see, e.g., Solanki et al., 2000; Stenflo, 2017; Mursula et al., 2007; Usoskin, 2017).

2.1. Short Time Scale Events

In the present forecast paradigm CMEs and flares are treated effectively as random events. Accuracy of timing an event depends on assessing which solar active regions are about to produce a CME/flare, and estimating when that will occur. The lead time to prepare for impact from such an event ranges from two to five days, depending on the speed with which the ejected material travels toward the Earth (for example, see Temmer et al., 2017, and references therein). This window allows for some preparatory mitigation steps.

The most plausible way for forecasting the onset of these events is probably by identifying “precursors,” which would cause an active region to erupt as a CME or flare and produce space weather events somewhat before they are observed to start on the Sun. It was pointed out by van Driel-Gesztelyi et al. (2002), using SoHO/MDI data, that CMEs are preceded by magnetic evolution during which helicity of the source region increases due to twisted flux emergence. This helicity increase can be a precursor to CMEs. From Hinode, observational estimates of outflow velocities of a mature active region inside an equatorial coronal hole showed intensified of such flows in a few days, during the complex interaction between the active region and coronal hole, led to a CME erupting from this active region (Baker et al., 2012). The generation of intensified flows can be explained by three-dimensional (3-D) MHD numerical simulations of Murray et al. (2010) who showed that compression of the nearby CH field by the expanding AR leads to outflows in field rooted in a same-polarity configuration. An AR observed by Baker et al. (2012), using Hinode EIS, exhibited a sigmoidal shape, indicating that a flux rope was present there. The slow rise and expansion of that flux rope prior to the eruption would, therefore, logically lead to stronger compression of the surrounding field and intensification of outflows. Thus, the intensification of outflows at the edge of this AR can be a precursor of its CME. The preeruption mechanisms and conditions for CMEs and flares are still a major area of ongoing research, which could lead to better predictions for short-time scale events.
2.2. Solar Seasonal and Subseasonal Time Scales (6–18 Months)

Ever since Hale’s (1908) discovery that sunspots contain strong magnetic fields (Hale, 1908) and the Maunders’ demonstration (see, e.g., Maund, 1905a, 1905b) that there was a significant association between solar and geomagnetic activity, it has been recognized that the Sun can strongly influence the interplanetary environment of the Earth. In the present days the term space weather has been used to identify these variations felt at the Earth. In the past few decades, it has been shown further that solar activity shows short-term variability with about 150–160 day periodicity, called the Rieger-type periodicity (Rieger et al., 1984), as well as with a quasi-biennial periodicity (Zaqarashvili et al., 2010b). Much more recently, McIntosh et al. (2015) showed that episodes of extreme solar activity occur as a result of global bursts in solar magnetism that take place on time scales of 6–18 months. Figure 2 shows multiple synchronized magnetic signatures of these bursts. These bursts, and the associated energetic output, have almost the same order-of-magnitude amplitude as the better-known decadal-scale variability known as the solar cycle. Thus, the eruptions that stimulate the most extreme CMEs and flares result in the bursty “seasons” of space weather.

Figure 2 demonstrates the appearance of these seasons for sunspot cycle 24. The five frames, to be read from bottom to top, depict manifestations of seasons first in emerging magnetic flux (bottom frame) and then in EUV bright point density (next frame up), M- and X-class solar flares (third frame), the daily flare rate (fourth frame), and finally, for comparison, a running average of the hemispheric and total sunspot number (top frame). Notice the strong periodicity present, with peaks corresponding to substantial increases (approaching 100% modulation) in sunspot and flare production (see the figures and discussion in McIntosh et al. (2015)).

Notice also the strong correspondence of the largest flares with the peaks of those bursts of magnetism. One of these peaks in cycle 24, which occurred in July 2012 in the south hemisphere of the Sun (see the blue curve in the top frame of Figure 2), led to the Carrington-type space weather event that just missed hitting the Earth. Despite the fact that the cycle 24 has been the weakest in 100 years, we still had such a major event. The original Carrington space weather event occurred in a similar bursty phase of cycle 10, which was also a relatively weak cycle. The main point we make from Figure 2 is that almost all energetic flares (X-class and M-class), which are characteristic of Carrington-type events, occur during bursty seasons.

Not only the energetic flares but also the CME rates strongly correlate with the bursty solar activity seasons. CMEs and solar energetic particle eruptions are intimately connected, and these phenomena are central to severe space weather impacts across many societally critical infrastructures. Higher frequencies of CMEs, together with heightened concentration of solar energetic particles, penetrate into the Earth’s atmosphere and reach the Earth’s magnetic field during bursts of solar activity. Although the solar flares occurring from the Carrington events in 1859 as well as in 2012 were associated with huge CMEs, the launches of major CMEs are distinct from the occurrence of large X-ray/EUV flares. There have been intensive discussions (see, e.g., Harra et al., 2016) that nine out of 42 solar flares, observed during February 2011 to November 2014 from the SDO and other sources, were not associated with CMEs. Tachocline Rossby waves we will be discussing here are particularly promising for determining bursty seasons of space weather. This is because CMEs are more closely linked to large-scale magnetic Rossby wave patterns at the surface and in the corona than the much smaller and more convoluted topological structures that induce flares (see, e.g., Török & Kliem, 2005) without any associated CMEs.

The large-scale organization in time and location (latitude-longitude) of eruptive events, which originate from the enhanced bursts of solar active regions, indicates their origin to be deep inside the solar interior, most likely in the tachocline (a shear layer), located at the interface of radiative core and convective envelope. The solar tachocline is very likely to contain strong magnetic fields. Seasonal events have been recently demonstrated to be intimately connected with solar tachocline Rossby waves’ dynamics (Dikpati et al., 2017; Dikpati, Belucz, et al., 2018; Dikpati, McIntosh, et al., 2018), which undergo complex interactions with tachocline differential rotation and spot-producing magnetic fields through quasi-periodic exchange of energies to drive TNOs.

Accurate simulations of their details will increase our ability to forecast the timing, amplitude, and locations of individual events. This is analogous to accurate forecasting of the combination of Rossby waves and jet streams in the Earth’s atmosphere that leads to better forecasting of individual large cyclonic storms, such as the forecast of Hurricane Sandy, which caused enormous damage in the Northeast United States in October–November 2012. Sandy was a successful prediction in terms of the strength of the event and its...
Figure 2. Global-scale surges of magnetic flux emergence manifest themselves in many forms and drive the most violent of solar storms over the relatively benign sunspot cycle 24 (2010–present). From bottom to top we display the daily evolution of the Sun’s magnetic field at the central meridian, where persistent concentrations of magnetic flux are visible; more visible even in the density of EUV bright points driven by small bipolar regions on the disk, coinciding with the red, densest, clusters of bright point production; considering energetic flare locations over the sunspot cycle (C-class in green, M-class in blue, and X-class in red) we see that the C-class flares grossly mimic the bright point (BP) density pattern while, more often than not, M- and X-class flares occur ONLY during flux emergence surges; taking all flares greater than C5 in magnitude (with a running 50-day average) in each solar hemisphere (red = north, blue = south) we see the bursty, quasi-periodic, nature of flaring activity and notice that the biggest, X-class, flares occur in surges of enhanced activity; using the largest-scale manifestation of solar magnetism we present the daily hemispheric sunspot number (with a running 50-day average) and again note the relationship between X-class flares and surges of activity. The top three panels use dashed lines to demonstrate the connection of X-class flares in the scheme of the flaring rate and sunspot production rates in each solar hemisphere. Clearly, these largest flares—those with most disruptive potential—do not occur at random and are very highly correlated with significant upturns in magnetic flux emergence. The fourth frame is adapted from McIntosh et al. (2015).
location. The successful prediction was based on accurate simulations of trains of Rossby waves interacting with the jet stream, including Rossby wave “breaking,” or wrapping up into intense extratropical cyclones (see, e.g., Pantillon et al., 2015). An analogous study for North Pacific tropical cyclone was carried out by Riboldi et al. (2019). Similarly, combinations of Rossby waves, differential rotation, and magnetic fields in the solar tachocline have the potential to predict X-class flares as part of a major eruption of new magnetic flux on the Sun during a bursty phase.

3. Concept of Rossby Waves

In 1939 Carl Gustav Rossby showed (Rossby, 1939), with a simple theory, that rotating thin spherical shells, such as the Earth’s atmosphere, should contain global longitudinally propagating waves of vorticity. Haurwitz, 1941 extended the theory to a full spherical shell. These Rossby waves, as they have come to be called, are a fundamental property of rotating systems that extend over a substantial latitude range. In their simplest form, Rossby waves are properties of thin rotating spherical shells of fluid, in which the fluid flow is purely horizontal, contained within the shell. The only forces at work are Coriolis forces, coming from the fact that the shell is rotating, and fluid pressure forces. The Coriolis forces are proportional to the component of the rotation vector that is perpendicular to shell, and so are proportional to the sine of the latitude. Accordingly, for a given velocity, Coriolis forces are largest at the poles of the rotating shell, and zero at the equator.

All oscillations in fluids must have a restoring force in order to oscillate. Solar physicists are very familiar with acoustic waves, for which the restoring force is the pressure gradient force. Similarly, gravity waves oscillate as a result of the buoyancy force, which is itself a difference of two forces, namely, the vertical pressure gradient force and the force of gravity. For Rossby waves, the restoring force is the local difference between the Coriolis and horizontal pressure gradient forces.

From the beginning of Rossby wave theory, meteorologists have preferred to describe Rossby waves in terms of vorticity of the flow, in mathematical terms the curl of the horizontal velocities in the spherical shell. In this formulation, the pressure forces are hidden, because mathematically they are the gradient of a scalar, namely, the fluid pressure. Since the velocities are strictly within the shell and parallel to its lower and upper boundaries, the vorticity vector points in the local vertical direction. Rossby and Haurwitz proved that the total vorticity at a point on the sphere is conserved following a fluid element as it moves around. This means that the sum of the vorticity of the flow relative to the rotating frame, and the local vorticity of the frame itself, given by $2 \times \Omega \times \sin(\text{latitude})$, must be conserved. This property lends itself to a graphical description of the Rossby wave mechanism, which we give in Figure 3.

In the northern hemisphere, fluid elements moving poleward acquire a clockwise spin and thereby move back southward, until they spin counterclockwise relative to the rotating frame and start to move poleward again. Figure 3 schematically describes the operation of Rossby waves in a thin, rotating spherical shell as a consequence of total vorticity conservation. The Rossby wave mechanism works because in hydrodynamic case there is no way to transfer relative vorticity from one fluid element, as shown in Figure 3, to another. The addition of magnetic fields creates a mechanism for transferring relative vorticity from one fluid element to another, which we discuss later in this subsection.

An individual Rossby wave propagates westward relative to the mean zonal flow; that is, the phase velocity of Rossby waves is retrograde. But the group velocity of Rossby waves is prograde, implying that a Rossby “wave train” propagates eastward (Regev et al., 2016). This is because Rossby waves are dispersive waves. If they were nondispersive, the group velocity and phase velocity will be the same. What we have described above is for Rossby waves in strictly two-dimensional flow on a spherical shell. In reality, in most systems, the flow is not strictly 2D, but rather contains also some small vertical motions, which have the same horizontal spatial scale as the horizontal flow. The shallow water system we will rely on extensively in this article is one such system. In this case, it is not the vorticity that is conserved, but rather a related quantity called the “potential vorticity” that is conserved. In the shallow water system, potential vorticity is the total vorticity divided by the shell thickness, which can also vary with location because the top boundary can deform. In even more general systems, with continuous vertical, subadiabatic stratification, a slightly different form of potential vorticity is conserved. Each of these more general systems also contain Rossby waves, with slightly modified properties.
The flow in Rossby waves is virtually horizontal in the spherical shell, and has a horizontal spatial scale much larger than the thickness of the layer, implying that the flow is hydrostatic. This generally requires that the vertical temperature gradient be “subadiabatic,” which further guarantees that the flow will be nearly horizontal. The term subadiabatic means that the specific entropy of the thin layer increases with radius, in contrast to the “superadiabatic” convection zone, where, on average, the specific entropy must decrease with radius, which provides potential energy to drive convection through buoyancy forces. In the subadiabatic layer, vertical displacements of fluid are resisted by negative buoyancy. In addition, the time scale for Rossby waves is relatively long compared to the rotation period. Other shorter-period waves can also be present, but these are not Rossby waves. On Earth, Rossby waves originate from the differential heating by the Sun, but that is not the only way they can be sustained.

Figure 3. (a) A schematic diagram shows how in a rotating shell of fluid Rossby waves can be generated when the fluid is perturbed in the north-south direction. The undisturbed vorticity at latitude \( \theta_0 \) is taken to be zero relative to the rotating reference frame. The leftmost and rightmost solid circles of fluid move poleward and the middle circle, equatorward. The open circles signify the initial locations of the displaced fluid. Due to conservation of total vorticity, the leftmost and rightmost fluid elements acquire anticyclonic swirl, that is, clockwise in the north hemisphere. (b) The middle fluid element acquires cyclonic swirl. The cyclonic and anticyclonic swirls are not confined to the magenta fluid elements, but influence the neighboring fluid elements in the way depicted by black curved arrows in (b). (c) As a result, the two undisplaced elements (second and fourth fluid elements) in (b) will move in latitude in the directions shown with heavy black arrows. In addition, because of vorticity conservation, the first, third, and fifth fluid elements reach their maximum latitudinal displacements and reverse direction in latitude, as depicted by the heavy black vertical arrows attached to them in (c). The result is a westward moving wave pattern as shown by the relative positions of the heavy black wave. The whole figure is modified from Figure 1 in Dikpati, McIntosh, et al. (2018).
Starting from the fundamental concept of Rossby waves, enormous progress has been made over the past 80 years in being able to forecast global weather patterns that are made up largely of Rossby waves coupled with upper level jet streams of longitudinal flow (see, e.g., www.weather.gov/jetstream).

All rotating celestial objects with fluid atmospheres and/or interiors have the potential to generate and maintain Rossby waves. In the Sun, there are multiple depths where Rossby waves might occur, but the most likely place is in the tachocline, where we know there is strong differential rotation and relatively weak or no convection. As with the Earth's troposphere where Rossby waves are found, the tachocline vertical temperature gradient is subadiabatic, again favoring nearly hydrostatic flow. By contrast, in the convection zone the strong convective motions imply nonhydrostatic flows, and flows whose horizontal scale is not large compared to the vertical scale. There might be some signs of Rossby wave-like behavior, such as longitudinal phase propagation, but other Rossby wave-like behavior may be obscured or even overpowered by the convection there. That said, so far, Rossby wave properties can be detected only in the solar atmosphere and a short distance down into the convection zone (McIntosh et al., 2017; Krista et al., 2018; Hanasoge & Mandal, 2019; Loptien et al., 2018; Liang et al., 2019). We are not yet able to determine the depth in the Sun where Rossby waves “originate.”

By definition, the tachocline separates the differentially rotating convection zone above from the solidly rotating radiative interior. The tachocline differential rotation is maintained by turbulent momentum transport down from above. Rossby waves can arise from instability of this differential rotation in latitude, which is the energy source for their growth. In this process, the differential rotation in latitude weakens, but is subsequently replenished by the convection zone above. Thus, Rossby waves in the tachocline have a different physical origin from those in the Earth’s atmosphere, but they are Rossby waves nonetheless. As the total vorticity is conserved for the Rossby waves in the Earth’s atmosphere, an analogous argument can be made for the solar tachocline Rossby waves that the so-called potential vorticity (vorticity divided by the thickness of tachocline fluid shell) is conserved.

That said, Rossby waves in the tachocline will differ from those on Earth in at least one very important way—they will be strongly influenced and modified by the presence of strong toroidal magnetic fields there, for which there is no counterpart in the Earth’s lower atmosphere. The solar plasma is highly electrically conducting, so any flow across field lines will immediately generate electric currents that alter the Rossby wave properties. Magnetic fields, particularly the spot-producing toroidal fields, fundamentally modify HD Rossby waves. These modifications depend on the amplitude and the profile of the magnetic fields present. In effect, they become MHD Rossby waves, in which the potential vorticity is no longer conserved; the fluid elements in different locations, such as depicted in Figure 3, are connected by magnetic fields that can propagate vorticity from one element to another by means of Alfven waves.

Any model of global dynamics of the tachocline must take account of these MHD effects. But it is precisely these MHD effects that have the potential to be linked to magnetic features and their evolution on the solar surface, which in turn can provide a link between global MHD of the solar interior, and space weather in the heliosphere. We discuss this in detail in the next section.

Beyond the Earth and Sun, MHD Rossby waves are very likely to occur in rotating stars with convection zones and therefore tachoclines, possibly in neutron stars and even in accretion disks around stars. The atmospheres and oceans of exoplanets, especially those that are not tidally “locked” with their host star, will surely also contain hydrodynamic Rossby waves. Planets with molten or liquid metal cores are also very likely to have a variety of MHD Rossby waves. MHD Rossby waves in strong magnetic fields in the Earth’s outer core have been studied theoretically for at least the 50 years. They may explain the origin of the westward drift of the geomagnetic field (Hori et al., 2015). Dynamo models of the core also exhibit Rossby wave patterns (Hori et al., 2018). Thus, Rossby waves, both HD and MHD, are a common feature of the global dynamics of rotating stars and planets.

4. Observations of Rossby Waves in the Sun

4.1. Early Indications of Short-Term Periodicity Due to Rossby Waves

Starting with Rieger et al. (1984), evidence has steadily accumulated of solar activity variations with periods a little less than six months. Rieger et al. found that “hard” (gamma-ray) solar flares tended to appear in
groups separated by about 154 days. Oliver et al. (1998) used wavelet analysis applied to sunspot areas to detect a 158-day periodicity that essentially coincided with that for the most energetic flares. This detection was expanded to include a 160-day period in photospheric magnetic flux during solar cycle 21 (Ballester et al., 2002), fading out during cycle 22, but reappearing in cycle 23 (Ballester et al., 2004). Javaraiah and Komm (1999) found peaks in sunspot rotation data at 2/3-, 1.2-, and 2.4-year periods. These early measures of semiannual periods were not at first clearly identified with Rossby waves, but Wolff (1992) provided theoretical evidence that they might be related to global oscillation modes of the rotating Sun. Dimitropoulou et al. (2008) used a model for such modes to make a clearer connection to Rossby waves. Since then the theory of Rossby waves for the Sun has developed rapidly, including toroidal magnetic fields, and become connected with the theory of linear and nonlinear global MHD waves and instabilities of the solar tachocline, all of which we discuss in later sections. Very recently, the theory of MHD Rossby waves has been used to estimate the strength of the toroidal field in the solar interior (Gurgenashvili et al., 2016), even distinguishing toroidal field amplitudes between north and south hemispheres (Gurgenashvili et al., 2017).

All of these studies have been focused on periods of oscillation, without measures of phase propagation in longitude, which is an essential feature of Rossby waves. To measure this phase propagation, making time-longitude diagrams is essential. This need leads to so-called Hovmöller diagrams, first used in meteorology.

4.2. Hovmöller Diagrams and Wave Patterns

Figure 2 captures evidence of quasi-periodically driven surges of magnetic flux, resulting in unstable magnetic configurations, which in turn give rise to the quasi-periodic occurrence of flares (and CMEs) on the principal magnetic activity bands, as cataloged by McIntosh et al. (2015). In addition, at latitudes higher than the principal activity bands, in the first few years of the descending phase of cycles, recurrent high-speed solar streams are observed (e.g., Legrand & Simon, 1981). The source of these fast wind streams appears to quasi-rigidly rotate in solar location—somewhat of a mystery if coronal holes are considered to result from active region decay. McIntosh et al. (2015) also demonstrated the quasi-periodic nature of the fast solar wind speed and density. These phenomena are not disconnected—they belong to magnetic activity bands of the present and upcoming solar cycles that are stacked in latitude, and present on the solar disk almost all of the time (McIntosh et al., 2014).

An essential diagnostic in the connection of the observed phenomena and the underlying process comes from the construction of a Hovmöller diagram (Hovmöller, 1949)—a common means of plotting meteorological data—used to show the propagation of pressure and rainfall patterns in longitude. The axes of a Hovmöller diagram are longitude (abscissa) and time (ordinate) for fixed latitudes. Slanting straight lines in a Hovmöller diagram are evidence of a sequence of disturbances propagating in longitude. Essentially Hovmöller diagrams are space-time plots. The left frame of Figure 4 depicts a circular band of information, representing the complete evolution of the system in longitude. The right frame shows the evolution of the pattern from a narrow range of latitudes.

The following two subsections use EUV images in the 195- and 193-Å broadband channels taken by the STEREO/EUVI and SDO/AIA telescopes, respectively, that are combined to construct synchonic images from which Hovmöller diagrams can be constructed. The principal emissions in these two instruments are subtly different, but the images carry similar information. Observing with three spacecraft enabled imaging of the entire solar atmosphere from 6 February 2011 until 1 July 2014.

Figure 5 depicts the STEREO Behind 195Å (in Figure 5a), SDO AIA 193Å (in Figure 5b), and the STEREO Ahead 195Å (in Figure 5c) images on 22 August 2011, around midnight (UT). Their combination is shown in Figure 5d. We map the three solar images onto heliographic coordinate grids—the longitude of the image centers was prescribed by the longitude of each spacecraft. The images are normalized relative to their pixel intensity histograms and superimposed onto a common grid with a prescribed pixel scale (0.1°), and we simply adopt a mean value for the composite image where there is overlap between the constituent images.

Figure 6 shows sample Hovmöller diagrams that can be constructed from the combined SDO/STEREO data set by averaging over 4° latitude, using the method described above. BP, like all features, are sparse, and so we do need to bin over time (and space) to accumulate “signals.” The imaging data do not have the same constraint. Therefore, daily samples of Hovmöller diagrams of BP are constructed using a 28-day running
average of instantaneous BP density observed by three spacecraft—SDO, STEREO-A, and STEREO-B. From left to right they show the evolution of coronal features at southern hemisphere midlatitudes, the southern hemisphere activity band, the northern activity band, and northern midlatitudes. Several features are clear in these plots, taking the midlatitudes first—we can clearly see the presence of coronal holes (elongated dark features) that slant to the left in both hemispheres, and some bright emission features, likely small active regions that are also slanted left. Since the Carrington rotation rate is defined at 38° latitude, the slant is not reflecting the local rotation rate, which should give nearly vertical or slightly prograde slanting features at the 35° (i.e., the leftmost and the rightmost frames in Figure 6). Recalling from above, and referring to the analysis presented in McIntosh et al. (2017), that those slants equate to an apparent retrograde motion of 55 m/s.

This retrograde motion must be due to something besides differential rotation and meridional circulation. Therefore, a global wave pattern, which is most likely the Rossby wave pattern, cause those dark features move retrograde. In contrast, the coronal holes present on the principal activity bands exhibit a mix of slants to the left and to the right, indicating a mixture of prograde and retrograde motions; note also that the gradients are steeper, indicating a slower apparent motion. These patterns are moving eastward more slowly than the local rotation at low latitudes; thus, the most likely cause of these motions is again Rossby waves. Section 3.2 discusses the consequences of the coronal hole evolution a little more. The activity band data show bright features that slant right individually but appear to be in groups, or clusters, that slant to the left; as we will see, in the following subsection, this is a determining characteristic of Rossby waves in a rotating system.

Attractive as this interpretation is, we must nevertheless keep in mind that different rotation rates might be generated by different magnetic structures, even sunspots of different sizes and lifetimes, that might be reflecting rotations at different depths (see, Javaraiah & Komm, 1999, and references cited therein). On the other hand, Snodgrass (1983) found no statistically significant time changes in surface rotation measured from magnetograms for the period 1967–1982.

Figure 4. Illustrating the concept of a Hovmöller diagram. (left) For a synchronous spherical data set a Hovmöller diagram is constructed by sampling the longitudinal evolution of the system in a narrow range of latitudes over time. (right) The annuli are then stacked to represent the passage of time. In the example shown the density of synchronous EUV bright point (BP) density is shown and was the original example used to diagnose the presence of Rossby waves in the global-scale magnetic flux systems of solar interior.
4.3. Rossby Waves in Coronal Bright Points

Several features are already clear from Figure 7, including the clear asymmetry of magnetic activity between the southern (Figure 7c) and northern (Figure 7a) hemispheres. Also, in the February–March 2011 time frame, there appears to be an abrupt switch-on of activity in the northern hemisphere that over time begins to encompass more longitudes. In the time frame shown, the southern hemisphere significantly lags the north, but short episodes of activity are noticeable.

Coronal BPs permit the tracking of the magnetic activity bands of the 22-year magnetic cycle of the Sun (McIntosh et al., 2014). As we have mentioned earlier, these activity bands in each solar hemisphere undergo significant quasi-annual instability which results in episodes of intensified space weather (McIntosh et al., 2015). McIntosh et al. (2017) used a BP detection algorithm on a series of coronal images taken by twin STEREO spacecraft and the SDO spacecraft and combined the detected BPs (at 4-hr cadence) across the solar sphere to construct latitude-longitude images of the BP density from which Hovmöller diagrams could then be constructed (see, e.g., the right panel of Figure 5). Typically, due to their sparsity, these BP diagnostic maps are running averages over time and over a range of latitudes. Figure 6 shows typical examples for the principal activity bands in the rising phase of sunspot cycle 24. Figure 7 shows the same key features as
Figure 6—minus the tilt of the ARs to the right. Note that these diagrams use a much wider band for averaging (15° compared to 4° latitude range used for Figure 6). Consequently, the tilt we see in Figure 6 gets somewhat obscured in Figure 7. This figure is to illustrate longitudinal behavior in the activity bands as the sunspot cycle is initialized. Note the large periods of no activity before the slanted features start in both hemispheres. The work this figure describes can be found in McIntosh et al. (2019).

We can clearly again relate surges in magnetic flux emergence, through modulation in the hemispheric sunspot number and multiple longitude activity of BPs—consider the period around the end of 2011 and start of 2012 where both hemispheres react. McIntosh et al. (2017) used the characteristics of the clusters of BP activity to capture the presence of two, oppositely directed velocities in the data as we have mentioned above; a slow ~3 m/s westward phase velocity, of individual clusters while analyzing the apparent motion of clusters, revealed a 24 m/s eastward group velocity—the classical signature of meteorological Rossby waves. Furthermore, McIntosh et al. (2017) noted that the BP clusters exhibit lifetimes that are integer numbers of the rotational period at the latitudes sampled, hinting further that the processes governing the magnetic flux emergence underlying the signal were intrinsically tied to the rotation of the plasma.
Figure 7. Coronal bright point (BP) density Hovmöller diagrams, derived from the joint observations of STEREO-A, STEREO-B, and SDO, show longitudinal evolution of BP in the (a) northern, (b) equatorial region, and (c) southern activity bands with respect to the (d) daily hemispheric sunspot number. The BP density is constructed using 28-day running average of instantaneous BP density of three spacecraft. These BP density Hovmöller diagrams illustrate features of rising phase of sunspot cycle 24 (onset of cycle 24 is indicated by dashed horizontal line). Gray shaded box in bottom panel indicates the time span (vertical axis) in the bottom panels.
4.4. Rossby Waves From Coronal Patterns

Using daily Hovmöller diagrams constructed from the EUV images of STEREO and SDO (see, e.g., Figure 7) the lifetimes and propagation characteristics of coronal holes in longitude were studied over several solar rotations (Krista et al., 2018). Three distinct populations of “low-latitude” (below 65° latitude) coronal holes were readily detected. One population rotates in retrograde direction and coincides with a group of long-lived (over 60 days) coronal holes in each hemisphere. These recurrent (from the perspective of the Sun-Earth line) coronal holes were located between 30° and 55°, and displayed velocities of around 55 m/s, slower than the local (differential) rotation rate (see Figures 6a and 6d). A second, smaller population of long-lived coronal holes were observed inside 10° latitude—those exhibit prograde motion with velocities between 20 and 45 m/s. A third population of CHs exist that are short-lived (they live less than two solar rotations). They appear over a very wide range of latitudes and also exhibit a wide range of velocities (−140 and 80 m/s). A “butterfly diagram” of the rising phase of sunspot cycle 24 showed a systematic evolution of the longer-lived holes at midlatitudes, appearing to connect them with the overlapping band picture developed by McIntosh et al. (2014) and hinted at by Legrand and Simon (1981), but the sample was too short to draw concrete conclusions.

As for the BPs explored by McIntosh et al. (2017), the rotational modulation is evident in the analysis of coronal hole lifetimes presented by Krista et al. (2018). Because the convective forces near the photosphere and just below would erode coronal hole structures within one rotation, it is likely that these patterns have roots deeper in the solar interior. Clearly, looking at coronal holes on the rotational time scale in a broader data set (e.g., Gibson et al., 2017) could yield additional properties of the Rossby modes present in the solar interior. Furthermore, given the role of coronal holes, in shaping the solar wind that permeates the heliosphere and on space weather throughout, understanding the longitudinal properties and/or characteristic lifetimes over many solar cycles at relevant epochs may go a very long way toward helping forecast the more stealthy variety of space weather and possibly close the loop on the investigations of Legrand and Simon (1981) and others.

4.5. Rossby Waves From Helioseismic Measurements

In addition to the evidence for solar Rossby waves from solar atmosphere data (described immediately above), there is also very recently accumulating evidence from helioseismic analyses (Hanasoge & Mandal, 2019; Liang et al., 2019; Loptien et al., 2018). They used time-distance, ring-diagram, and mode coupling techniques to extract information from varying lengths of SDO/HMI measurements, most recently Liang et al. (2019) analyzing 21 years of such data. By the nature of helioseismic data, the results are confined to low and middle latitudes (<40° or so). Loptien et al. (2018) focused specifically on what they call equatorial Rossby waves. These estimates of Rossby waves are also limited to the outermost few percent of the solar radius, namely, a thin layer near the top of the convection zone.

The overall results are in general agreement, with some differences. The principal result is that all three studies found evidence for westward (retrograde) propagating Rossby waves whose frequency relation with longitudinal wave number, m, follows rather well the long-known theoretical relationship (originally derived in Haurwitz, 1941) for waves in a rotating thin spherical shell in the absence of any differential rotation. A spectrum of longitudinal wave numbers out to m = 15 or so was detected. There was also found evidence that the flow patterns in the waves showed substantially more vorticity about a vertical axis than they did horizontal divergence, another well-known property of Rossby waves. Lifetimes of individual waves were found to be within a few months and a year or somewhat more.

All these measurements match best with Rossby waves that are not influenced by magnetic fields, for which we know (Dikpati, Belucz, et al., 2018) that the phase speed is significantly changed in the tachocline by the presence of spot-producing toroidal magnetic fields. Their retrograde speeds are also not always the same as found from coronal measurements, but we need to keep in mind that those measurements are tied to properties of the Sun’s magnetic field. It is possible the helioseismic and coronal measures are detecting different waves. The difference in Rossby wave speeds from coronal and helioseismic measures is most pronounced near the equator.

We focus later on models of nonlinear MHD Rossby waves in the solar tachocline, which is far below where the helioseismic Rossby wave detections have been made. It is certainly possible that there are separate...
Over the past two decades many quasi-3-D shallow water and 2-D tachocline models have been built and explored for understanding the various waves, instabilities, and nonlinear interactions present in the system.

### 5. Models for HD and MHD Rossby Waves in the Sun

In the years since the original paper on Rossby waves (Rossby, 1939), there has been enormous development of both linear and nonlinear hydrodynamic models for these waves in differentially rotating thin spherical shell. One of the most useful models is the so-called “shallow water” model (see, e.g., Pedlosky, 1987). Since there exist several books already which describe the history and development of hydrodynamic shallow water models, applied to the oceanic and atmospheric systems, we do not discuss the development of such models in detail here.

The most commonly used shallow water model is one that has a single spherical shell of fluid, for which the top surface is allowed to deform. The fluid is incompressible and in hydrostatic balance in the vertical, even though vertical motions are allowed. Velocities well-simulated with shallow water models are those of global dimensions in longitude and latitude in layers that are thin compared to the radius of the rotating body. Shallow water models are well-suited for studying Rossby waves (see, e.g., Enagonio & Montgomery, 2001) without the full complexity of 3-D hydrodynamics. Shallow water systems are normally inviscid, representing “ideal fluids.” Shallow water systems conserve the volume-integrated total energy (kinetic and potential) for the entire fluid shell. The one-layer shallow water system includes strictly two-dimensional modes as a limiting case for which the subadiabatic stratification (or effective gravity; see next paragraph) becomes very large.

We apply a shallow water model to the solar tachocline, which is a thin layer at the base of the convection zone, where the differential rotation of the convection zone above declines to the solid rotation of the deep solar interior. We identify that the center of the tachocline is roughly located at the base of the convection zone (~0.7 R); thus, the bottom half of our shallow water tachocline model is located near the outer boundary of the radiative interior and the top half is in the overshooting layer of the convection zone. Because the convection zone is nearly adiabatically stratified, it yields easily to material rising from below. Shallow water models are generally characterized by a nondimensional parameter, called the “effective gravity” ($G_{eff}$), which is proportional to the fraction by which the stratification is subadiabatic. Since the subadiabaticity is large in the radiative zone, but much smaller in the overshoot zone, $G_{eff}$ of the part of the tachocline in the radiative interior is high, and in the overshoot part it is low. Since the tachocline is very likely to contain strong magnetic fields, adding magnetic fields to such a model is essential for studying Rossby waves in the Sun. Gilman (2000) generalized shallow water equations to include MHD.

Even though the top boundary deforms, magnetic fields remain confined to the layer. No field lines cross either boundary, and fields tangent to the boundary remain there always. Vertical components of velocity and magnetic field are linear functions of depth, whereas the horizontal components are independent of depth. The total pressure, including fluid and magnetic, is hydrostatic, and its gradient is proportional to the horizontal gradient of shell thickness. There are no magnetic monopoles, because the total magnetic flux is conserved. This is captured by a divergence-free field condition: the horizontal divergence of the product of the height and horizontal field is zero. The total energy including kinetic, potential, and magnetic energies is conserved.

Simulation models based upon the physical foundations and approximations described above are capable of generating multiple types of global waves that have characteristics of Rossby, gravity, and Alfven waves. With the presence of differential rotation and latitudinally varying toroidal magnetic fields, these waves can be either neutral or unstably growing forms. Nonlinear interactions among the waves and also with the ambient differential rotation and toroidal fields will be common. There will be substantial exchanges of energy of various forms (potential, kinetic, and magnetic) among the mean flows and fields and waves. Therefore, simulation models must be robust and general enough to capture all of these physical effects. Over the past two decades many quasi-3-D shallow water and 2-D tachocline models have been built and explored for understanding the various waves, instabilities, and nonlinear interactions present in the system.
5.1. Neutral Waves

To understand the physics of the MHD shallow water system outlined above, it is very helpful to linearize the governing equations, and find the linear modes of oscillation as well as exponentially growing unstable modes. Oscillatory modes come from restoring forces in the system, which in the solar tachocline include at least Coriolis, buoyancy, and electromagnetic body forces (leaving aside acoustic modes, which are in a much higher-frequency range and do not contribute to global dynamics of the whole system). These forces are what allow, respectively, Rossby, gravity, and Alfven waves in the system. Since all these restoring forces are present together, the actual waves will arise from various combinations of forces. Of greatest significance here are magnetically modified Rossby waves. In the presence of differential rotation and toroidal fields that vary in latitude and radius, these waves can become unstable and grow exponentially. But they can retain some propagation properties of neutral waves, particularly Rossby waves. The energy sources for the exponential growth are found in the energy stored in the differential rotation and toroidal field, and even in the variable thickness of the shallow water tachocline. The paths the energy can follow were diagrammed in Gilman and Fox (1997) and Dikpati and Gilman (2001).

Over the past decade, there has been extensive theoretical analysis of the various Rossby-type wave modes that could be present in the solar tachocline. The models used have most often been based on the MHD shallow water system, beginning with Zaqarashvili et al. (2007), who showed the existence of both “fast” and “slow” MHD Rossby waves, the former being essentially hydrodynamic Rossby waves slightly modified by the field, and the latter, the result of strong fields that lead to magnetic body forces that work against Coriolis forces, leading to the low frequency. The analysis has been extended (Zaqarashvili et al., 2009) to include the stable stratification of the tachocline, and showed a class of waves that are trapped near the solar equator. Zaqarashvili et al. (2010a, 2010b) have further identified the modes of lowest longitudinal wave number with both Rieger-type periodicities and quasi-biennial oscillations, as well as to variations in magnetic activity over much longer time periods—hundreds of years (Zaqarashvili et al., 2015). Most recently (Zaqarashvili, 2018), a special form of equatorial “magneto-Kelvin” wave has been identified with the 11-year sunspot cycle, and fast magneto-Rossby-gravity and magneto-inertia-gravity waves have been associated with Rieger-type periodicities. In parallel studies, Schecter et al. (2001); Raphaldini and Raup (2015); and Klimachkov and Petrosyan (2017a, 2017b) have looked at nonlinear interactions among these shallow water waves.

While these comparisons of linear wave frequencies to observed frequencies found in solar observations are suggestive of a physical link, it is necessary to distinguish between frequency as a measure of longitude propagation speed, as discussed in the previous paragraph, and frequencies with origins in amplitude oscillations, such as the sunspot cycle, and solar seasons. The relations between these two types of frequencies is a continuing subject of research.

5.2. Unstable Waves

In the presence of differential rotation and toroidal fields, Rossby, gravity, and Alfven waves can all be present. Extensive modeling shows that for expected conditions in the tachocline, only magnetically modified Rossby waves will interact with the differential rotation and toroidal fields in such a way as to make the background state unstable, allowing exponential growth of these waves (Arlt et al., 2005; Cally, 2003; Cally et al., 2008; Dikpati et al., 2003; Dikpati & Gilman, 1999, 2005; Gilman, 2000; Gilman, 2015, 2017; Gilman & Dikpati, 2000; Gilman & Fox, 1997). By contrast both gravity waves and Alfven waves, particularly for strong fields, have frequencies and phase velocities that are much too high to allow energetically active interaction with the differential rotation. In a similar way acoustic waves in the Sun will not be unstable drawing energy from the differential rotation.

In terms of kinematics, since Rossby waves propagate in longitude at a rate that is close to or within the range of longitudinal velocity of the differential rotation, the streamlines appear almost stationary in a rotating frame (Gilman & Fox, 1997). This allows the differential rotation to create “tilts” with latitude in the streamlines. By tilt, we mean that the maximum poleward location of a particular streamline of a Rossby wave is either further to the west or to the east of the adjacent streamline at a lower latitude. The perturbed toroidal fields will also have tilts. Both of these tilts imply velocity and magnetic stresses that can transport momentum in latitude. These transports, in turn, will lead to conversions of energy between differential
rotation and toroidal fields, on the one hand, and perturbation-kinetic and perturbation-magnetic energies in the Rossby waves, on the other. By perturbation energy we mean the departure from energy of the unperturbed reference state. We discuss the role of tilted streamlines and field patterns in section 5.4. By contrast, gravity and Alfvén wave patterns are generally not tilted, because they propagate too fast to be affected much by the differential rotation.

There are multiple paths of energy flow that can allow instability to develop. The dominant path is primarily determined by the relative amounts of energy contained in the profiles of reference state differential rotation and toroidal fields. In qualitative terms, if the background toroidal fields are weak, unstable modes will be driven primarily by kinetic energy stored in the differential rotation. For stronger background toroidal fields, particularly when toroidal field energy exceeds that of the differential rotation, both will be energy sources for instability, with the toroidal field energy source becoming dominant as toroidal field is increased. But some differential rotation must be present to facilitate the instability. In all cases, for linear, exponentially growing modes, these energy sources are effectively “infinite.” This is because, in linear models, there are no feedback from the growing perturbations to modify them to account for the extraction of energy.

In a hydrodynamic model of solar tachocline the eigenfunctions of exponentially growing Rossby modes, plotted as a planform map in latitude and longitude, show eastward tilts with latitude in the perturbation streamlines (see, e.g., Dikpati & Gilman, 2001, Figure 7). But the perturbation streamlines and field lines show westward tilts with latitude in an MHD shallow water tachocline, when magnetic fields dominate the dynamics (see, e.g., Gilman & Dikpati, 2002). In the hydrodynamic regime the Reynolds stress transports angular momentum toward the poles, whereas in the magnetohydrodynamic regime Maxwell stress does the same, to drive the instability. In both regimes the disturbances are Rossby waves, whose structure and longitudinal propagation vary according to the strength and location of the toroidal fields. As mentioned in the previous paragraph, the energy reservoirs are infinite sources to drive the instability in the linear case; therefore, the unstable modes maintain their tilts. When nonlinear effects are included, the tilts are modified by differential rotation as the Rossby wave amplitude grows, which ultimately shuts off energy extraction from the reference state reservoirs, and even reverses the energy flow back toward building up the differential rotation and/or toroidal field.

Both energetically neutral and unstable growing MHD Rossby waves have phase speeds in longitude. In general, there are far more neutral oscillatory Rossby waves than there are unstable ones. So there are potentially far more comparisons that can be made between phase speeds of neutral modes and observations of Rossby waves, such as those discussed in section 4. On the other hand, unstable modes are bound to reach finite amplitude due to their self-excitation, while neutral modes require some other excitation mechanism. If we can show that the longitudinal phase speeds of even linear unstable Rossby waves agree with observed speeds, this supports the idea that the observed Rossby waves are the unstable modes. Dikpati, Belucz, et al. (2018) have shown that the most unstable MHD Rossby waves in tachocline differential rotation for banded toroidal fields have approximately the same speed as the rotation rate at the latitude of the toroidal band. This agrees well with the surface measures of phase speed from coronal holes and bright points at the same latitude. This suggests that the origin of the measured phase speed is unstable MHD Rossby waves. It also suggests that Rossby wave phase speeds can give an estimate of the latitude location of the tachocline toroidal field at any given time. Dikpati, Belucz, et al. (2018) point out that these Rossby waves are the same ones responsible for “solar seasons,” as described in section 6.

5.3. Nonlinear MHD of Rossby Waves

In the nonlinear evolution of a nearly dissipationless system of tachocline latitudinal differential rotation, magnetic fields, and Rossby waves, an oscillation among energies of longitude-averaged flow and fields, and Rossby waves, should be expected. Dikpati (2012) built the first fully nonlinear hydrodynamic shallow water tachocline model for Rossby waves interacting with differential rotation. Dikpati et al. (2017) and Dikpati, McIntosh, et al. (2018) recently found such back-and-forth exchange of energy among the reference state differential rotation, magnetic fields, and the Rossby waves in tachocline. These oscillations are the so-called TNOs, which are demonstrated to have profound influence in governing the quasi-periodic seasons of space weather.
Figure 8 schematically displays the basic physical mechanism behind the simplest type of TNO, in which the oscillatory exchange occurs between tachocline differential rotation and Rossby waves. In principle Rossby wave streamlines can be tilted either westward or eastward with latitude in the presence of the differential rotation, but only the ones tilting eastward will grow, because that sense of tilt allows them to extract kinetic energy from differential rotation via the Reynolds stress. Thus, the kinetic energy of differential rotation decreases in this phase. Subsequently, the differential rotation is no longer able to feed energy to the Rossby waves in this nonlinear phase; as a result, Rossby waves reach a peak in amplitude. The waves start losing their eastward tilt, eventually tilting westward and feeding energy back into the mean flow. Thus, a nonlinear oscillation is created between the Rossby wave and differential rotation, very much like the nonlinear Orr mechanism (Orr, 1907) in fluid mechanics.

The TNO also manifests itself in oscillatory bulges and depressions of the top boundary of the shallow water tachocline. The tachocline toroidal magnetic flux is most likely to escape the tachocline in those places in latitude and longitude where the tachocline experiences maximum upward swelling (Dikpati & Gilman, 2005). When the Rossby wave reaches its peak amplitude, the tachocline top surface is maximally deformed. This is the time when the eruption of magnetic flux from the tachocline top surface to the photosphere is maximized, producing a bursty season. This bursty phase then declines producing the quieter “season” when the differential rotation amplitude is restored. But it is also possible for the nonlinear MHD Rossby
wave model to produce shorter, more random periods of activity, simply from coincidence of spot-producing toroidal fields and tachocline top surface bulges lasting for shorter intervals of time.

5.4. Simulations of Rossby Wave Interactions With Differential Rotation

A quasi-periodic exchange of energy, occurring in a typical simulation between the reference state differential rotation and the Rossby wave disturbances, is shown in Figure 9. In Figure 9a the kinetic energy of differential rotation ($K$) has been plotted in solid red, and the kinetic energy of Rossby wave-like perturbations ($K'$) in dashed red. Rossby wave undergoes an antiphase oscillation with the differential rotation; that is, $K'$ is at its maximum when $K$ is minimum and vice versa. We mathematically prescribe the differential rotation ($\omega$), as $\omega = s_0 - s_d \mu^2 - s_d \mu^4 - \omega_c$ (\(\mu\) is sine latitude); in the case presented in Figure 9, the total differential rotation is 21% of the interior rotation ($\omega_c$). In this case, $s_2/s_0 = 0.15, s_d/s_0 = 0.06$ and the effective gravity $G_{eff} = 0.5$. Initial amplitude of perturbation Rossby waves is ~40%, which means that the initial perturbation kinetic energy ($K'$) is 16% with respect to the kinetic energy of differential rotation ($K$). There is also potential energy in the shallow water system. ($P$) is the energy associated with longitude-averaged deformations, and $P'$

Figure 9. (a) Tachocline nonlinear oscillations (TNOs) are shown. Solid and dashed red curves show the oscillation between differential rotation energy ($K$); solid red curve) and Rossby waves' kinetic energy ($K'$). The x and y axes show dimensionless units, in which 100 time units correspond to one year. Simulations for these TNOs included ~42 modes in longitude. In this case, the TNO period is approximately six months. (b–d) Arrow vectors in the right panel show Rossby wave patterns, and color shades represent tachocline fluid shell's thickness (red denotes swelling and blue depression). In top frame (b) of right panel, due to eastward tilts, Rossby wave patterns extract energy from differential rotation until Rossby waves' energy is at a maximum, and differential energy cannot supply energy any longer; thus, Rossby wave patterns lose their tilts, go through neutral tilts (c), and then they tilt westward (d). (e) Enhanced activity bursts occur when Rossby wave energy reaches maximum (this epoch is shown by a semitransparent gray arrow, which is pointing to a local peak; “bursty season” marked by solid yellow ellipse—the sunspot number curve). Bursty season is followed by a relatively quieter season (marked by the second solid yellow ellipse in the sunspot curve; figure adapted from Dikpati et al. (2017)).
is the energy arising from the longitude average of the local potential energy associated with longitude-dependent deformations. $\overline{\mathcal{P}}$ and $\mathcal{P}'$ are, respectively, shown in solid and dashed blue curves. The black curve traces the evolution of the total energy $(\mathcal{K} + \mathcal{K}' + \overline{\mathcal{P}} + \mathcal{P}')$, which is nicely conserved.

Figures 9b–9d show the time evolution of disturbance patterns. In Figure 9b the Rossby waves have maximum eastward tilts with latitude to extract maximum energy from differential rotation. Subsequently, in Figure 9c, we can see that $\mathcal{K}'$ has increased and $\overline{\mathcal{K}}$ has decreased, so the epoch is reached when the tilts are neutral and no more energy is extracted from $\mathcal{K}$. The tilts then become westward (see Figure 9d), allowing energy transfer from the Rossby waves to differential rotation. As a result, the differential rotation is rebuilt, and the process repeats, leading to the tachocline nonlinear oscillations.

The example shown in Figure 9 produced a TNO of about six-month periodicity for the selected differential rotation amplitude. The periodicity of the bursts is a function of differential rotation amplitude as well as the effective gravity. Figure 5 of Dikpati, McIntosh, et al. (2018) shows a more complex case of TNO, for which the period evolves with time, making the TNO more quasi-periodic.

When the amplitude of the Rossby waves is the largest, the top surface of the tachocline is most deformed, allowing tachocline magnetic fields to enter the convection zone and make their buoyant rise to the surface. Thus, a bursty season of flux eruption events occur, followed by a decline in Rossby wave amplitude, top surface deformation, and hence a relatively quieter season of bursts. The TNO period in this model determines the periodicity of enhanced bursts of activity, and hence the seasons of space weather.

### 5.5. MHD Rossby Waves and TNOs

The tachocline is most likely to contain spot-producing toroidal fields, which is the east-west field, a magnetic counterpart of the east-west flow (i.e., the differential rotation in the Sun). Both HD and MHD Rossby waves are expected to occur, and participate in the interaction with the reference state differential rotation and magnetic fields. The HD and MHD Rossby wave patterns (which work as velocity and magnetic perturbations) tend to have similar tilt patterns in longitude and latitude, but with a significant longitudinal phase difference between them. When both the HD and MHD Rossby wave patterns are tilted “upstream” away from the equator (very much like the red ellipses in Figure 8c), there is angular momentum transport toward the equator by the Reynolds stress in both hemispheres. But the Maxwell stress from the upstream-tilted field lines signifies angular momentum transport away from the equator.

If the Maxwell stress is larger than the Reynolds stress, then the Rossby waves will grow by extracting energy from differential rotation. Even if the toroidal field is weak, Rossby waves can grow by extracting energy from the east-west flow (the differential rotation) (see, e.g., Gilman & Fox, 1997). The oscillation “starts” by the Maxwell stress extracting energy from the east-west flow (the differential rotation) at a greater rate than the Reynolds stress can replenish it. The former is transporting angular momentum toward the poles, and the latter toward the equator, as shown near the top of Figure 8. The energy extracted goes into the perturbation velocities and magnetic fields, in the form of MHD Rossby waves. This has the effect of weakening the differential rotation. But this implies that there is less kinetic energy still available to extract, and as a result, the Rossby wave amplitude ceases to grow. The perturbation streamlines and field lines are no longer tilted. But the system does not stop evolving there; instead, the tilts reverse, and energy comes out of the Rossby waves and back into the differential rotation.

When the toroidal field energy is large compared to differential rotation energy, we still get the same sequence of tilts, but the primary source of energy for growing Rossby waves becomes the toroidal field, with differential rotation acting more as a catalyst to facilitate the energy exchange. Getting energy out of the toroidal field requires a phase shift in longitude between velocity and magnetic perturbations. The stronger the field is, the larger this phase shift is. In this case the nonlinear growth of MHD Rossby waves is halted when the toroidal field is weakened enough by energy extraction that it can no longer supply further growth. Then the phase difference between fields and velocities lessens, and the toroidal field can rebuild.

Cally et al. (2003) found that in purely 2D (no free surface deformations), nonlinear MHD Rossby wave simulations produced primarily a simple “tipping” of toroidal field about an equilibrium east-west orientation, so all perturbation energy was in modes with longitudinal wave number $m = 1$. The total magnetic energy of the ring remained constant, with the tipping with time measuring the energy of wave numbers.
m = 0 and m = 1 oscillating in antiphase. But in typical 3-D case, we see deformation of the toroidal fields. At any given time, the m = 0 toroidal field has less energy in it than it began with, to supply energy to the perturbations, so it must be wider in latitude at all later times. As the nonlinear oscillations proceed, the phase difference in longitude between velocity and magnetic fields widens during perturbation growth and then shortens, perhaps even changing sign, during perturbation decline. At the same time, perturbation field lines reverse their tilt when the differential rotation reaches its minimum, then allowing it to rebuild.

For a “strong field” case, for which the energy in the toroidal field is large compared to that in the differential rotation, Rossby waves still grow, but their energy source is primarily from the toroidal field, rather than from the differential rotation. From the strong toroidal field, there is a strong $\mathbf{j} \times \mathbf{B}$ perturbation body force that does work to create perturbation kinetic energy, which is quickly converted into even stronger perturbation magnetic energy, characteristic of an MHD Rossby wave, that is rather different from its purely hydrodynamic counterpart. In these MHD Rossby waves, tachocline bulges and depressions in the top surface are still excited, and can play their role in providing longitude and latitude locations for emergence of toroidal fields into the convection zone and photosphere, but the potential energy associated with these deformations is small compared to the magnetic and even kinetic energy reservoirs.

It is important to recognize that the amplitude oscillations described above are totally distinct from the longitudinal phase speeds and frequencies of MHD Rossby waves, even nonlinear ones. This is true even though the period of a longitudinally propagating Rossby wave (the time it takes for a particular phase in the wave to travel 360° around a circumference) may be close to that of period of the amplitude oscillation. Longitudinal phase propagation is a linear phenomenon, while the amplitude oscillation of a TNO is strictly nonlinear. The restoring forces that determine the frequencies in the two cases are completely different. Thus, an MHD Rossby wave whose phase takes many years to propagate along a circumference (and so has a very long period) cannot explain an amplitude oscillation of the same period.

5.6. Energy Reservoirs and Conversions

We have described above how Reynolds, Maxwell, and mixed stresses of MHD Rossby waves can extract energy from the differential rotation and toroidal fields to excite and maintain their own amplitudes. These energy conversions can be described pictorially with a so-called energy diagram, which depicts all the energy reservoirs and the energy conversion processes that connect them. The details of such an energy diagram are derived and presented in details in Dikpati, McIntosh, et al. (2018), particularly in the Appendix. All such energy conversions involve work done by the fluid against the forces present, including inertial, pressure, and $\mathbf{j} \times \mathbf{B}$ body forces. Therefore, there are energy conversion links between potential and kinetic energy reservoirs, and between kinetic and magnetic energy reservoirs, but there are no links directly between toroidal field energy and magnetic energy of the Rossby waves. It is straightforward to prove that the total energy (potential plus kinetic plus magnetic) is conserved. This is a key property that leads to the system being capable of nonlinear oscillations, since there are no dissipative “leaks,” such as would be present for a damped oscillator like a pendulum swinging in a viscous medium.

5.7. Robustness of TNO Periods

The actual period of oscillation in the nonlinear MHD shallow water model for the solar tachocline is determined by the physical inputs to the model: the differential rotation, effective gravity (parameter G), and the amplitude and profile with latitude of the toroidal field present. Dikpati et al. (2017) have studied how the period varies with these physical parameters, showing that the period declines slowly with an increase in differential rotation amplitude. The period also declines from about 20 months for overshoot values of G to as low as two months for the radiative part of the tachocline. In general, for given differential rotation and effective gravity, the oscillation period shortens with increasing toroidal fields, and for the same toroidal band, decreases with increasing latitude of band placement. But for all these conditions, the TNO period remains in a range consistent with observational evidence. In qualitative terms, for weak magnetic fields (say, <5 kG), the oscillations are hydrodynamic even though toroidal field is present. But for strong toroidal fields, the oscillations are strongly influenced by the electromagnetic body force.

It is to be noted that Gurgenashvili et al. (2016) already showed that Rieger-type periods of magnetic Rossby waves are shorter during stronger cycles and longer during weaker cycles. It is reasonable to assume that stronger cycles originate from stronger tachocline toroidal fields. The TNO period shortens for increasing...
magnetic field band amplitude (see, e.g., Figures 4c and 4d in Dikpati et al. (2017), for details). Therefore, observed periodicity can be a useful measure of spot-producing toroidal magnetic field strength for different cycles.

In a very recent theoretical study Gachechiladze et al. (2019) explored the dynamics of magneto-Kelvin and magneto-Rossby waves by employing an MHD shallow water model of the solar tachocline; they showed that the global magnetic Rossby waves are primarily responsible for producing the observed quasi-annual oscillations in the solar activity.

6. Progress in Forecasting Space Weather by Data Assimilation

One important aspect of space weather forecasting is to estimate the time of major eruptive events; the other is to forecast where they occur on the Sun—only some events have the Earth in the “bullseye.” The timing and frequency of the major events, which have been observationally demonstrated to occur during the bursty seasons, are primarily determined by the TNOs. Estimating the locations (latitude-longitude) and intensity of these bursts is also a goal and should be possible.

Magnetogram observations indicate that the surface locations of magnetic flux eruptions are not random; instead, they follow certain patterns. Figure 10 shows that the latitude-longitude of solar active regions, often called “active longitudes,” are organized in time (Benevolenskaya et al., 1999; de Toma et al., 2000) in such a way as to reveal a large-scale longitudinal wave pattern for several Carrington rotations. The longitudinal mode with wave number 1 persists for several Carrington rotations, then includes wave number 2 and higher as the time of solar activity progresses.

Due to the persistence of active longitudes for several Carrington rotations, the Earth feels strongly periodic variations in solar output just because the Sun rotates. If there is just one active longitude, then this period is essentially the same as the Sun’s rotation period. But at any given time, more than one active longitude along a full circumference (see, e.g., the top frames in Figure 10) can lead to shorter periods, which would differ according to the longitude spacing between the different active longitudes. Hence, the space weather seasons are also linked with active longitudes’ locations and strengths.

The existence of active longitudes on the Sun has been recognized at least as far back in time as Maunder (1905b), and continues to be an active area of observational analysis. A particularly important question is how long they persist. There are estimates of lifetimes ranging from several solar rotations up to as many as 10 sunspot cycles (Benevolenskaya et al., 1999; Bumba et al., 2000; Usoskin et al., 2007; Obridko, 2009; Hathaway, 2015; Mandal et al., 2017) as well as evidence of both uniform and differential rotation of active longitudes (Berdyugina et al., 2006; Plyusnina, 2010). Families of MHD Rossby waves in the solar tachocline have strong potential for producing such persistence of longitude-dependent features. Individual waves will propagate in longitude with a phase velocity that is independent of latitude, while at the same time a family of Rossby waves can persist for long time periods without rapid dispersion in longitude.

Such a well-organized evolution of active longitudes again indicates their deep origin. Shallow water TNO models provide a plausible physical origin of the formation and evolution of active longitudes. These models show persistent patterns of bulges and depressions of tachocline fluid with low longitudinal wave numbers. Portions of the solar dynamo-generated, sunspot-producing, toroidal magnetic bands that coincide with bulging of the fluid are likely to emerge buoyantly to the surface through the convection zone (see Figure 11). This bulging leads to magnetic flux emergence and the enhanced formation of active regions. Note that observations indicate that occasionally new active regions appear in a more random way, not subsequently identified with an active longitude. Rossby waves interacting with the mean flow and mean toroidal (spot-producing) fields in a TNO model must simulate both kinds, namely, the regular and systematic eruptions in the form of active longitudes, as well as irregular and random eruptions.

We take the concept illustrated in Figure 11 a step further in Figure 12, in which we illustrate how tachocline bulges in toroidal field can be matched with surface magnetograms showing emerged active regions, for three consecutive recent Carrington rotations (CRs), 1923, 1924, and 1925 (CR start dates of, respectively, ~22 May, ~18 June, and ~15 July 1997). The yellow arrows in each hemisphere connect bulges of toroidal field in our MHD tachocline model with the observations. Our purpose here is not to give a detailed case...
Figure 10. (left) Synoptic maps of magnetic flux in northern hemisphere (0°–90°) of the Sun during 1996–1998. (right) The same in the southern hemisphere. In left (right) panel, the solar equator is at the bottom (top) (figure adapted from de Toma et al. (2000) with permission).
study from simulations, but rather to simply show how this matching could work in practice. Note that by the third CR the match has started to degrade, because a new active region has emerged that does not match well with a tachocline bulge. The red arrow from tachocline to photosphere illustrates this mismatch. Such a mismatch can be postponed and reduced by using data assimilation applied to surface magnetic data to update the model simulations.

Two major classes of data assimilation techniques have been used for several decades in atmospheric and oceanic applications, namely, the Ensemble Kalman Filter methods, and variational methods. The Ensemble Kalman Filter method is a statistical method involving sequential evolution of state space with bias caused by observations, and the variational method depends on the evolution of models in adjoint space to recreate known observations, by creating an adjoint model either directly or by numerical differentiation. In the solar context, application of data assimilation is still in its infancy, but is growing fast. Ensemble Kalman Filter data assimilation has been applied in the geodynamo and various solar models by several authors (Arge et al., 2010; Dikpati et al., 2014; Dikpati, Anderson, & Mitra, 2016; Dikpati, Mitra, & Anderson, 2016; Fournier et al., 2010; Hickman et al., 2015; Kitiashvili, 2018), and also, the variational approaches have been incorporated for estimating parameters in solar dynamo models (Jouve et al., 2011; Hung et al., 2017).

To compare our solar model outputs with observations, and to integrate the model forward for forecasting the evolution of magnetic fields, their eruption timing, locations, strengths, and sizes, it is necessary to assimilate the available observations.

By assimilating observations of active regions from magnetograms (primarily their sizes, strengths, and locations) at every few hours to few days, we update the initial conditions of spot-producing magnetic fields at the tachocline level, and more accurately simulate the evolution of active region locations as well as their periodic bursts at the surface.

7. Utility and Value of Forecasting Rossby Waves

We have discussed in great detail how Rossby waves, interacting with the Sun’s interior shear fluid layer, can determine the timing of quasi-periodic seasonal bursts of solar activity that shape space weather in longitude and time. As we have discussed above giant meandering terrestrial Rossby waves are associated with the global atmospheric pressure environment and jet stream through which they determine the large-scale tropospheric weather patterns. Through physical analogy and the exploitation of analytic and numerical techniques...
developed in terrestrial meteorology we feel that tracking the Sun's Rossby waves, their nonlinear dynamics, and their longitudinal evolution will result in a considerable increase in our ability to forecast changes in the space environment. We anticipate that the wealth of observational data (including diagnostics of Rossby waves and magnetogram observations for activity bursts) married with data assimilation techniques and recently developed nonlinear MHD models described above will permit the accurate prediction of the magnetic activity bursts well in advance of energetic event onset. This improved lead time should significantly reduce the risks that the Sun, and space weather, pose to technologically dependent society.

**Figure 12.** Sequence of three successive Carrington rotations (CR) showing patterns of simulated spot-producing fields (white ribbons) and bulges (red) and depression (blue) of the tachocline top surface. Yellow arrows depict locations and upward tracks of magnetic flux that has originated from places where the toroidal field and tachocline bulges coincide. Observed surface magnetic activity is shown in gray scale at the top of each CR “box.” The red arrow in the bottom frame indicates where a new active region emerged that does not correspond to an upward bulge in the tachocline and its toroidal field.
Appendix A

As discussed in section 4.1, a shallow water model is a quasi-three-dimensional fluid system. In our case the fluid is the solar plasma in the tachocline. It includes the radial dimension in the tachocline fluid shell in a simplified way, by allowing the upper boundary to vary in latitude, longitude, and time, so that radial motions can occur and the thickness of the fluid shell can vary. The core approximation is that the horizontal scale of fluid and the disturbances in it are very large compared to the undisturbed thickness of the fluid, so the pressure perturbations are hydrostatic. Due to the thinness of the tachocline compared to solar radius the divergence of the radii and the density variation are ignored in the momentum and mass continuity equations.

The ideal MHD shallow water model equations can be written in the vector-invariant laboratory frame form (see Gilman, 2000) as follows:

\[
\frac{\partial B}{\partial t} + U \cdot \nabla B = B \cdot \nabla U - (1 + h) B = 0,
\]

\[
\frac{\partial h}{\partial t} + U \cdot \nabla h = h \cdot \nabla U,
\]

\[
\frac{\partial U}{\partial t} = \frac{1}{2} \nabla (B + U) - (\times U) \cdot \nabla x \frac{g H}{(1 + h)} + (\times B) \cdot \nabla x B,
\]

in which \(B\) is the horizontal magnetic field which is the function of horizontal coordinates and time; \(U\) is the horizontal velocity, the unit vector in the vertical direction; \(\nabla\) is the horizontal component of the curl operator; \(g\) is the gravity at the tachocline depth; and \(H(1 + h)\) is the thickness of the tachocline fluid shell where \(H\) is the undisturbed constant thickness. The evolution equation for the fractional layer thickness \(h\) is fundamentally derived from the mass continuity equation, and it is derived from the assumption that dynamics are nearly incompressible (an asymptotically true statement when the time scales are long compared to acoustic propagation times).

The full set of nonlinear MHD shallow water equations in nondimensional form in the rotating frame (rotating with the Sun’s core rotation rate) in spherical coordinate can be found in Dikpati, McIntosh, et al. (2018). The numerical algorithm and the technique for solving nonlinear shallow water equations for the solar tachocline in spherical polar coordinate in the rotating frame already exist (see, e.g., Dikpati, 2012, Appendix, for details; Dikpati, McIntosh, et al., 2018). Here we briefly discuss the major steps taken in solving the nonlinear MHD shallow water system. The height-deformation variable \(h\) is decomposed in scalar spherical harmonics and the horizontal velocity \((u, v)\), and magnetic field components \((a, b)\) (i.e., latitude-longitude components in spherical coordinate) in vector spherical harmonics to deal with the pole problem. Nonlinear terms in the equations are computed following pseudo-spectral implementation given in Swarztrauber (1996). Fourth order Runge-Kutta time integration scheme is implemented for time evolution; semiimplicit dynamics is included following Hack and Jacob (1992). This allows the use of larger time steps. Computer intensive synthesis and analysis steps are run concurrently in multiple parallel threads on modern many core processors.

The pseudo-spectral numerical algorithm with adaptive time-stepping implemented in our shallow water tachocline model is robust and therefore well-suited for simulating longer time scale phenomena, such as the dynamical interaction between magnetically modified Rossby waves and tachocline differential rotation, as well as much shorter time scale phenomena such as suddenly excited, steep gravity waves.

The force balance of the undisturbed system is given by the balance among three latitudinal forces: horizontal pressure gradient, magnetic curvature stress, and the Coriolis force. The model parameters include the “effective gravity” \(G\) of the tachocline, differential rotation amplitude, the strength and latitude location of toroidal magnetic field bands, and latitudinal width of these bands \(\alpha_c\) is core rotation rate, used as the rotating reference frame. This is approximately equal to the tachocline rotation rate at 32° latitude, the solar radius at the tachocline depth, and \(1/\alpha_c\) are, respectively, the dimensionless length and time. Thus, 100 dimensionless time units correspond to about one year. A complete description of the MHD shallow water equation parameters is given in Dikpati, McIntosh, et al. (2018).
Acknowledgments

We thank Ricky Egeland and Peter Gilman for internally reviewing this manuscript, and for their helpful discussions. We also extend our thanks to three anonymous reviewers for their thorough review of our previous manuscript, and for their valuable comments, which have helped improve our paper significantly. Data for Figure 2 are from SIDC/SILCO, GOES satellite, and SDO/HMI. Data can be available, respectively, from the websites http://sidc.oma.be/, https://www.goes.nws.noaa.gov/, and http://hmi.stanford.edu/. Data for Figure 7 are from NASA’s STEREO and SDO missions. STEREO data can be available from the website https://stereo-ssc.nasa.gov/data.shtml. Some of the other figures contain data are from previous publications and the reader can consult the appropriate citation. A major part of this material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under cooperative agreement 1852977, NASA’s LWS award 80NSSC20K0355, and a subaward from NASA’s HSR award 80NSSC18K1206 originally awarded to NSO.

References

Antia, H. M., Basu, S., & Chitre, S. M. (1998). Solar internal rotation rate and the latitudinal variation of the tachocline. *Monthly Notices of the Royal Astronomical Society*, 298(2), 543–556.

Arge, C. N., Henney, C., Koller, J., Compeau, C. R., Young, S., MacKenzie, D., et al. (2010). Air Force Data Assimilative Photopheric Flux Transport (ADAPT) model. In *Twelfth International Solar Wind Conference, AIP Conference Proceedings* (Vol. 1216, pp. 343–346).

Arul, R., Sulé, A., & Rudiger, G. (2005). Three-dimensional stability of the solar tachocline. *Astronomy & Astrophysics*, 441(3), 1171–1175.

Baker, D., van Driel-Gesztelyi, L., & Green, L. M. (2012). Forecasting a CME by spectropolarimetric precursor? *Solar Physics*, 276(1-2), 219–239.

Ballester, J. L., Oliver, R., & Carbonell, M. (2002). The near 160 day periodicity in the photospheric magnetic flux. *Astrophysical Journal*, 566(1), 505–511.

Ballester, J. L., Oliver, R., & Carbonell, M. (2004). Return of the near 160 day periodicity in the photospheric magnetic flux during solar cycle 23. *Astrophysical Journal Letters*, 619(2), L173-L176.

Benevolenskaya, E. E., Konovichev, A. G., & Scherrer, P. H. (1999). Structure and dynamics of interconnecting loops and coronal holes in active longitudes. *Solar Physics*, 190(1-2), 145–151.

Berdyugin, S., Moss, D., Sokolov, D., & Usoskin, I. G. (2006). Active longitudes, nonaxisymmetric dynamos and phase mixing. *Astronomy & Astrophysics*, 445(2), 703–714.

Bumba, V., Garcia, A., & Klvana, M. (2000). Longitudinal distribution of solar magnetic fields and activity during the ending and starting periods of activity cycles. *Solar Physics*, 196(2), 403–419.

Cally, P. S. (2003). Three-dimensional magnetosheath instabilities in the solar tachocline. Monthly Notices of the Royal Astronomical Society, 339(4), 957–972.

Cally, P. S., Dikpati, M., & Gilman, P. A. (2003). Clamshell and tipping instabilities in a two-dimensional magneto-hydrodynamic tachocline. *Astrophysical Journal*, 582(2), 1190–1205.

Cally, P. S., Dikpati, M., & Gilman, P. A. (2008). Three-dimensional magnetosheath instabilities in the solar tachocline: I—Axisymmetric case. *Monthly Notices of the Royal Astronomical Society*, 391(2), 891–900.

Canxon, P., Angling, M., Barclay, L., Curry, C., Dyer, C., Edwards, R., et al. (2013). Extreme space weather: Impacts on engineered systems and infrastructure. London: Royal Academy of Engineering.

de Toma, G., White, O. R., & Harvey, K. L. (2000). A picture of solar minimum and the onset of solar cycle 23. I. Global magnetic field evolution. *Astrophysical Journal*, 529(2), 1101–1114.

Dikpati, M. (2012). Nonlinear evolution of global hydrodynamic shallow-water instability in the solar tachocline. *Astrophysical Journal*, 745(2), 128.

Dikpati, M., Anderson, J. L., & Mitra, D. (2014). Ensemble Kalman filter data assimilation in a Babcock-Leighton solar dynamo model: An observation system simulation experiment for reconstructing meridional flow speed. *Geophysical Research Letters*, 41, 5361–5369. https://doi.org/10.1002/2014GL061777.

Dikpati, M., Anderson, J. L., & Mitra, D. (2016). Data assimilation in a solar dynamo model using Ensemble Kalman Filters: Sensitivity and robustness in reconstruction of meridional flow speed. *Astrophysical Journal*, 828(91), 18.

Dikpati, M., Belucz, B., Gilman, P. A., & McIntosh, S. W. (2018). Phase speed of magnetized Rossby waves that cause solar seasons. *Astrophysical Journal*, 86(2), 159.

Dikpati, M., Cally, P. S., McIntosh, S. W., & Heffez, E. (2017). The origin of the “seasons” in space weather. *Scientific Reports*, 7, Article number: 14750, pp 7.

Dikpati, M., & Gilman, P. A. (1999). Joint instability of latitudinal differential rotation and concentrated toroidal fields below the solar convection zone. *Astrophysical Journal*, 521(2), 417–441.

Dikpati, M., & Gilman, P. A. (2001). Analysis of hydrodynamic stability of solar tachocline latitudinal differential rotation using a shallow-water model. *Astrophysical Journal*, 551(1), 536–564.

Dikpati, M., & Gilman, P. A. (2005a). A shallow-water theory for the Sun’s active longitudes. *Astrophysical Journal Letters*, 635(2), L193-L196.

Dikpati, M., Gilman, P. A., & Rempel, M. (2003). Stability analysis of tachocline latitudinal differential rotation and coexisting toroidal band using a shallow-water model. *Astrophysical Journal*, 596(1), 680–697.

Dikpati, M., McIntosh, S. W., Bothun, G., Cally, P. S., Ghosh, S. S., Gilman, P. A., & Umurhan, O. M. (2018). Role of interaction between magnetic Rossby waves and tachocline differential rotation in producing solar seasons. *Astrophysical Journal*, 853(2), 144.

Dikpati, M., Mitra, D., & Anderson, J. L. (2016). Role of response time of a Babcock-Leighton solar dynamo model in meridional flow-speed reconstruction by Enhanced Kalman data assimilation. *Advances in Space Research*, 58(8), 1589–1595. https://doi.org/10.1016/j.asr.2016.08.004.

Dimitropoulou, M., Moussas, X., & Strintzi, D. (2008). Enhanced Rieger-type periodicities’ detection in X-ray solar flares and statistical validation of Rossby waves’ existence. *Monthly Notices of the Royal Astronomical Society*, 386(4), 2278–2284.

Enagionio, J., & Montgomery, M. T. (2001). Tropical cyclogenesis via convectively forced vortex Rossby waves in a shallow water primitive equation model. *Journal of the Atmospheric Sciences*, 58(7), 685–705.

Fournier, A., Hulet, G., Jault, D., Kiang, W., Tangborn, A., Gillet, N., et al. (2010). An introduction to data assimilation and predictability in geomagnetism. *Space Science Reviews*, 155, 247–291.

Gachchiladze, T., Zaqarashvili, T. V., Gugenashvili, E., Ramishvili, G., Carbonell, M., Oliver, R., & Ballester, J. L. (2019). Magneto-Rossby waves in the solar tachocline and the annual variations in solar activity. *Astrophysical Journal*, 874(2), 162.

Gibson, S. E., Webb, D., Hewins, I. M., McFadden, R. H., Emery, B. A., Denig, W., & McIntosh, P. S. (2017). Beyond sunspots: Studies using the McIntosh archive of global solar magnetic patterns. *IAU Symposium*, 328, 93.

Gilman, P. A. (2000). Magneto-hydrodynamic “shallow” waves equations for the solar tachocline. *Astrophysical Journal Letters*, 544(1), L79-L82.

Gilman, P. A. (2015). Effect of toroidal fields on baroclinic instability in the solar tachocline. *Astrophysical Journal*, 801(1), 22.

Gilman, P. A. (2017). Baroclinic instability in the solar tachocline for continuous vertical profiles of rotation, effective gravity, and toroidal field. *Astrophysical Journal*, 842(2), 130.

Gilman, P. A., & Dikpati, M. (2000). Joint instability of latitudinal differential rotation and concentrated toroidal fields below the solar convection zone. II. Instability of narrow bands at all latitudes. *Astrophysical Journal Letters*, 528(1), 552–572.

Gilman, P. A., & Dikpati, M. (2002). Analysis of instability of latitudinal differential rotation and toroidal field in the solar tachocline using a magneto-hydrodynamic shallow-water model. I. Instability for broad toroidal field profiles. *Astrophysical Journal*, 576(2), 1031–1047.
Gilman, P. A., & Fox, P. A. (1997). Joint instability of latitudinal differential rotation and toroidal magnetic fields below the solar convection zone. *Astrophysical Journal*, 484(1), 439–454.

Gurgunashvili, E., Zagorashvili, T. V., Kukhianidze, V., Oliver, R., Ballester, J. L., Dikpati, M., & McIntosh, S. W. (2017). North-south asymmetry in Rieger-type periodicity during solar cycles 19–23. *Astrophysical Journal*, 845(2), 137.

Gurgunashvili, E., Zagorashvili, T. V., Kukhianidze, V., Oliver, R., Ballester, J. L., Ramishvili, G., et al. (2016). Rieger-type periodicity during solar cycles 14–24: Estimation of dynamo magnetic field strength in the solar interior. *Astrophysical Journal*, 826(1), 55.

Hack, J. J., & Jacob, R. (1992). Description of a global shallow water model based on the spectral transform method. NCAR Tech. Note, NCAR/TN-343+STR.

Ihle, G. E. (1908). On the probable existence of a magnetic field in sun-spots. *Astrophysical Journal*, 28, 315.

Jouve, L., Brun, A. S., & Talagrand, O. (2011). Assimilating data into and atmosphere blocking downstream of recurving western North Pacific tropical cyclones. *Monthly Weather Review*, 147(2), 567–589.

Klimachkov, D., & Petrosyan, A. (2017b). Parametric instabilities in shallow water magnetohydrodynamics of astrophysical plasma in solar active regions. *Geophysical Research Letters*, 42, 6622–6629.

Klimachkov, D., & Petrovan, A. (2017a). Rossby waves in the magnetic fluid dynamics of a rotating plasma in the shallow-water approximation. *Journal of Experimental and Theoretical Physics*, 125(4), 597–612.

Klimachkov, D., & Petrovyan, A. (2017b). Parametric instabilities in shallow water magnetohydrodynamics of astrophysical plasma in external magnetic field. *Physics Letters A*, 381(2), 106–113.

Krista, L. D., McIntosh, S. W., & Leamon, R. Y. (2018). The longitudinal evolution of equatorial coronal holes. *Astronomy Journal*, 155(4), 153.

Legrand, J. P., & Simon, P. A. (1981). Ten cycles of solar and geomagnetic activity. *Journal of Experimental and Theoretical Physics*, A3.

Legrand, J. P., & Simon, P. A. (1981). Ten cycles of solar and geomagnetic activity. *Solar Physics*, 70(1), 173–195.

Liang, Z.-C., Gizon, L., Birch, A. C., & Duvall, T. L. (2019). Time-distance helioseismology of solar Rossby waves. *Astronomy & Astrophysics*, 626, A3.

Lopitien, B., Gizon, L., Birch, A. C., Schou, J., Prokaua, B., Duvall, T. L., et al. (2018). Global-scale equatorial Rossby waves as an essential component of solar internal dynamics. *Nature Astronomy*, 2(7), 568–573.

Mandels, S., Chatterjee, S., & Banerjee, D. (2017). Solar active longitudes from KodaiKanal white-light digitized data. *Astrophysical Journal*, 835(1), 62.

Mann, M. E. (2019). The Weather Amplifier: Strange waves in the jet stream foretell a future of heat waves and floods. *Scientific American*, 2019, 43.

Maunder, E. W. (1905a). The solar origin of terrestrial magnetic disturbances. *Astrophysical Journal*, 21, 101.

Maunder, E. W. (1905b). Distribution of sun-spots. *Monthly Notices of the Royal Astronomical Society*, 65, 556.

McIntosh, S. W., Leamon, R. J., Egeland, R., Dikpati, M., Fan, Y., & Rempel, M. (2019). What the sudden death of solar cycles can tell us about the nature of the solar interior. *Sol. Phys.*, 284, 7, article id. 88, 24 pp.

McIntosh, S. W., Leamon, R. J., Pichardo Marcano, M., & Leamon, R. J. (2017). The detection of Rossby-like waves on the Sun. *Nature Astronomy*, 1(4), 0086.

McIntosh, S. W., Leamon, R. J., Krista, L. D., Title, A. M., Hudson, H. S., Riley, P., et al. (2015). The solar magnetic activity band interaction and instabilities that shape quasi-periodic variability. *Nature Communications*, 6(1), 6491.

McIntosh, S. W., Wang, X., Leamon, R. J., Davey, A. R., Howe, R., Krista, L. D., et al. (2014). Deciphering solar magnetic activity. I. On the relationship between the sunspot cycle and the evolution of small magnetic features. *Astrophysical Journal*, 792(1), 12.

Murray, M. J., Baker, D., van Driel-Gesztelyi, L., & Sun, J. (2010). Outflows at the edges of an active region in a coronal hole: A signature of active region expansion? *Solar Physics*, 261(2), 253–269.

Mursula, K., Usoskin, I., & Maris, G. (2007). Introduction to space climate. *Advances in Space Research*, 40(7), 885–887.

Obridko, V. N. (2009). Large-scale patterns and “active longitudes”. Proceedings IAU Symposium No. 264, eds. A. G. Kosovichev, A. H. Andrei & J.-P. Rozelot.

Orik, V. M. (1907). Stability or instability of motions of a liquid. *PRDAA*, 27, 9.

Orik, V. M. (1907). Stability or instability of motions of a liquid. *PRDAA*, 27, 9.

Pantillon, F., Charbonneau, J.-P., & Richard, E. (2015). Time-periodicities of the Sun’s magnetic field strength in the solar interior. *Nature* (7), 568–573.

Plyusnina, L. A. (2010). Determination of the rotation periods of solar active longitudes. *Tellus*, 62(1), 439–449.

Raphaldini, B., & Raup, C. F. M. (2015). Nonlinear dynamics of magnetohydrodynamic Rossby waves and the cyclic nature of solar magnetic activity. *Astrophysical Journal*, 799(1), 78.

Regov, O., Umurhan, O. M., & Yecko, A. (2016). *Modern fluid dynamics for physics and astrophysics*. Graduate Texts in Physics. New York: Springer.
Rieger, E., Share, G. H., Forrest, D. J., Kanbach, G., Reppin, C., & Chupp, E. L. (1984). Rosby waves on the Sun as revealed by solar “hills”. Nature, 312(5995), 623–625.

Rosby, C.-G. (1939). Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. Journal of Marine Research, 2(1), 38–55.

Schechter, D. A., Boyd, J. F., & Gilman, P. A. (2001). “Shallow-water” magnetohydrodynamic waves in the solar tachocline. Astrophysical Journal Letters, 551, L85.

Schrijver, C. J., Kauristie, K., Aylward, A. D., Denardini, C. M., Gibson, S. E., Glover, A., et al. (2015). Understanding space weather to shield society: A global road map for 2015–2025 commissioned by COSPAR and ILWS. Advances in Space Research, 55(12), 2745–2807.

Snodgrass, H. B. (1983). Magnetic rotation of the solar photosphere. Astrophysical Journal, 270, 288–289.

Solanki, S. K., Schüssler, M., & Fligge, M. (2000). Evolution of the Sun’s large-scale magnetic field since George Ellery Hale. Space Science Reviews, 210(1-4), 5–35.

Temmer, M., Reiss, M. A., Nikolic, L., Hofmeister, S. J., & Veronig, A. M. (2017). Preconditioning of interplanetary space due to transient CME disturbances. Astrophysical Journal, 835, 141.

Török, T., & Kliem, B. (2005). Confined and ejective eruptions of kink-unstable flux ropes. Astrophysical Journal, 630, L97–L100.

Usoskin, I. G. (2017). A history of solar activity over millennia. Living Reviews in Solar Physics, 14(3), 3.

Usoskin, I. G., Berdyugina, S. V., Moss, D., & Sokoloff, D. D. (2007). Long-term persistence of solar active longitudes and its implications for the solar dynamo theory. Advances in Space Research, 40, 951–958.

van Driel-Gesztelyi, L., Schmieder, B., & Poedts, S. (2002). Magnetic build-up and precursors of CMEs. Proceedings of the Second Solar Cycle and Space Weather Euroconference, 24-29 September 2001, Vico Equense, Italy. Editor: Huguette Sawaya-Lacoste. ESA SP-477, Noordwijk: ESA Publications Division, ISBN 92-9092-749-6, 2002, p. 47–54.

Wolff, C. L. (1992). “Intermittent” solar periodicities. Solar Physics, 142(1), 187–195.

Zaqarashvili, T. V. (2018). Equatorial magnetohydrodynamic shallow water waves in the solar tachocline. Astrophysical Journal, 856(1), 32.

Zaqarashvili, T. V., Carbonell, M., Oliver, R., & Ballester, J. L. (2010a). Magnetic Rosby waves in the solar tachocline and Rieger-type periodicities. Astrophysical Journal, 709(2), 749–758.

Zaqarashvili, T. V., Carbonell, M., Oliver, R., & Ballester, J. L. (2010b). Quasi-biennial oscillations in the solar tachocline caused by magnetic Rosby wave instabilities. Astrophysical Journal Letters, 724(1), L95–L98.

Zaqarashvili, T. V., Oliver, R., & Ballester, J. L. (2009). Global shallow water magnetohydrodynamic waves in the solar tachocline. Astrophysical Journal Letters, 691(1), L41–L44.

Zaqarashvili, T. V., Oliver, R., Ballester, J. L., & Shergashvili, B. M. (2007). Rosby waves in “shallow water” magnetohydrodynamics. Astronomy & Astrophysics, 470(3), 815–820.

Zaqarashvili, T. V., Oliver, R., Hanslmayer, A., Carbonell, M., Ballester, J. L., Gachechiladze, T., & Usoskin, I. G. (2015). Long-term variation in the Sun’s activity caused by magnetic Rosby waves in the tachocline. Astrophysical Journal Letters, 805(2), L14. https://doi.org/10.1088/2041-8205/805/2/L14